

AFRL-RI-RS-TR-2010-075
Final Technical Report
March 2010



JESPP: JOINT EXPERIMENTATION ON SCALABLE PARALLEL PROCESSORS SUPERCOMPUTERS

University of Southern California

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RI-RS-TR-2010-075 HAS BEEN REVIEWED AND IS APPROVED FOR
PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/
STANLEY LIS
Work Unit Manager

/s/
MICHAEL J. HAYDUK, Acting Chief
Advanced Computing Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE*Form Approved*
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**1. REPORT DATE (DD-MM-YYYY)**

MARCH 2010

2. REPORT TYPE

Final

3. DATES COVERED (From - To)

June 2005 – September 2009

4. TITLE AND SUBTITLEJESPP: JOINT EXPERIMENTATION ON SCALABLE PARALLEL
PROCESSORS SUPERCOMPUTERS**5a. CONTRACT NUMBER**

N/A

5b. GRANT NUMBER

FA8750-05-2-0204

5c. PROGRAM ELEMENT NUMBER

N/A

6. AUTHOR(S)

Dan M. Davis

5d. PROJECT NUMBER

JESP

5e. TASK NUMBER

PS

5f. WORK UNIT NUMBER

05

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)Information Sciences Institute
University of Southern California
4676 Admiralty Way, Suite 1001
Marina del Rey, CA 90292-6695**8. PERFORMING ORGANIZATION
REPORT NUMBER**

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)AFRL/RITB
525 Brooks Road
Rome NY 13441-4505**10. SPONSOR/MONITOR'S ACRONYM(S)**

N/A

**11. SPONSORING/MONITORING
AGENCY REPORT NUMBER**
AFRL-RI-RS-TR-2010-075**12. DISTRIBUTION AVAILABILITY STATEMENT***APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED. THIS REPORT IS THE RESULT OF CONTRACTED
FUNDAMENTAL RESEARCH DEEMED EXEMPT FROM PUBLIC AFFAIRS SECURITY AND POLICY REVIEW.***13. SUPPLEMENTARY NOTES****14. ABSTRACT**

This research has been manifestly successful in advancing the state of the art in large-scale military battlespace simulations and in conceiving, developing, implementing and testing trans-continentially distributed high performance computing for training, analysis and evaluation. Significant achievements include increasing the scale of simulations by one order of magnitude, championing a 256-node General Purpose Graphics Processing Unit (GPGPU) cluster, facilitating compute environment stability, developing an effective Distributed Data Grid, enabling vital advances required to meet Joint Forces Command's (JFCOM's) mission goals and designing the JLogger system for improved logging. The research has resulted in the acceptance of 37 peer-reviewed research papers. In performing this research, ISI developed needed capabilities for JFCOM experiments, developed a fault tolerant router network, successfully proposed a 256-node GPGPU cluster, taught a course on its use, created the JLogger system, enabled faster analysis and investigated review techniques. All of this has had a significant salutary impact on the defense posture of the nation, resulting in direct benefits to the warfighter.

15. SUBJECT TERMS

High Performance Computing, Heterogeneous Computing, Distributed Data Management

16. SECURITY CLASSIFICATION OF:**17. LIMITATION OF
ABSTRACT****18. NUMBER
OF PAGES****19a. NAME OF RESPONSIBLE PERSON**
Stanley Lis**a. REPORT**
U**b. ABSTRACT**
U**c. THIS PAGE**
U

UU

87

19b. TELEPHONE NUMBER (Include area code)
N/A

Table of Contents

1.	Summary.....	1
2.	Tasks	2
3.	Problems.....	3
4.	Methods.....	5
5.	Assumptions	8
6.	Procedures.....	11
7.	Results and Discussion.....	18
8.	Conclusions.....	29
9.	Recommendations	32
10.	Management and Personnel Assigned	35
11.	References	36
12.	List of Acronyms.....	39
	Appendix A:.....	41
	Appendix B:	43

List of Figures

Figure 1: The Jfcom Clusters: Koa, Glenn And Joshua.....	11
Figure 2: Course In Gpgpu Programming - Jfcom, Oct 2007.....	13
Figure 3: Graph Of Entity Count Grow By Time On 14 Dec 07.....	18
Figure 4: Joshua	21
Figure 5: Splasch Block Diagram	23
Figure 6: Splasch Test Gui.....	24
Figure 7: Sdg's Cube Editor	27
Figure 8: Hpcs Computing Functionality Contributing To Fms.....	33

List of Tables

Table 1: Route-Planning Algorithm Classifications	15
---	----

Abstract

This research, conducted by the Information Sciences Institute (ISI) over the last four and a half years, has been manifestly successful in advancing the state of the art in large-scale military battlespace simulations and in conceiving, developing, implementing and testing trans-continentially distributed high performance computing for training, analysis and evaluation. Significant achievements include increasing the scale of simulations by one order of magnitude, championing a 256-node General Purpose Graphics Processing Unit-enhanced (GPGPU) cluster, facilitating compute environment stability, developing an effective Distributed Data Grid, enabling vital advances required to meet Joint Forces Command's (JFCOM's) mission goals and designing the JLogger system for improved logging. The quality of this research is verified by the acceptance of 37 peer-reviewed research papers. In performing this research, ISI developed needed capabilities to support Human-in-the-loop simulations for JFCOM experiments, further developed a trans-continental router network capability with a fault tolerant architecture, conceived and drafted a proposal for a 256-node GPGPU-enhanced cluster to meet JFCOM experiment needs, taught a course on its use, improved logging via the JLogger system, enabled faster analysis via better data management and investigated review techniques. All of this has had a significant salutary impact on the defense posture of the nation, resulting in direct benefits to the warfighter.

1. Summary

The Joint Forces Command (JFCOM) mission has been to lead the transformation of the United States Armed Forces to enable them to exert broad spectrum dominance as described in Joint Vision 2010 and 2020. In support of the mission of JFCOM, the Information Sciences Institute (ISI) of the University of Southern California (USC) has shown that High Performance Computing (HPC) is a necessary element in the computational tools required for that mission. The first step undertaken by ISI was the implementation of the Joint Semi-Automated Forces (JSAF) program on a PC-cluster, Scalable Parallel Processor (SPP) Supercomputer. This was accomplished in December of 2002 and made operationally useful in GFY 2003. In 2004, that capability was extended and expanded to make the SPPs an increasingly stable and useful tool in an operational environment. In addition, ISI provided operations support personnel at the Space and Naval Warfare Systems Command (SPAWAR) San Diego, California, Topographic Engineering Center (TEC), Ft. Belvoir, Virginia and JFCOM, Suffolk, Virginia.

ISI has “fielded” over ten million entities to meet Joint Urban Operations (JUO) experiment requirements. (Lucas, 2003; Lucas, 2009) All of this was made possible by using large Linux clusters. The clusters were funded and provided by the High Performance Computing Modernization Program (HPCMP) of the Department of Defense. These were made available to JFCOM following the proposals, which were drafted by the ISI team, for a Distributed Center for JFCOM’s use. ISI then further assisted in providing liaison with HPCMP, enabling their installation and utilization. This is a computational power otherwise not available to U.S. military planners and experimenters. The Information Sciences Institute (ISI) of the University of Southern California (USC) accomplished more than anticipated, utilizing their scalable computing technology and expertise.

That new increment of effort was required to enable future operations for military experimenters by ensuring the continuing availability of SPP’s and providing additional enhancements and extensions that were critically needed. The inherent scalability engendered in the ISI design allowed sufficient computing power to be applied to all of the required areas.

The efforts set forth above allowed a steadily decreasing level of effort with respect to the utilization by Joint Experimentation of the SPP platforms, while it facilitated research and development on data management capabilities. Those changes were reflected in the assignment of new personnel and new tasking for existing JESPP personnel.

2. Tasks

Therefore, several tasks were accomplished in GFYs 05(Qtrs. 3&4), 06, 07, 08 and 09:

- Developed capabilities and supported JUO and other JFCOM experiments
- Developed and implemented a secure, effective nation-wide router network capability for HPC operations (with an informal goal of supporting two million entities.)
- Investigated a fault tolerant architecture
- Conceived, but did not implement fault tolerant operations
- Investigated upcoming HPCMP DC compute requirements
- Developed training and documentation in support of SPP programs and procedures
- Improved the existing logging process to be faster, lighter and higher capacity (Giga Bytes (GB) to Tera Bytes)
- Investigated techniques for improving the analysis capability for much quicker results
- Investigated near real time Future After Action Review (FAARS) implementation
- In addition, JFCOM and AFRL added new tasking that was performed during the research.
- Proposed and facilitated HPCMP-provided, 256 node, GPGPU-enhanced cluster
- Conceived, designed, developed, and initially implemented a working instantiation of the Distributed or Scalable Data Grid
- Conceived, designed, developed, and implemented the JLogger System

3. Problems

The major thrust of all of this work was the advancement of the research capabilities for Human-In-The-Loop (HITL) battlespace simulations. These are required by JFCOM in order to conduct experiments using Agent Based Modeling (ABM) to look into the future of the U.S. defense structure. The research issues revolve around the need to simulate the participants in the 21st Century battlespace in ways that are more valid, of higher resolution, larger numbers of agents, on more extensive terrain, and exhibit more sophisticated behaviors. Additionally, faster execution times and better visualization are required. Failure to achieve these improvements will significantly constrain the advances required to ensure the security of our way of life.

In order to make productive use of the human participants, the individual workstations and the HPC Linux cluster, a secure, effective nation-wide router network is necessary, such need being exacerbated by the need to keep latencies low enough to not disrupt the simulation nor to discombobulate the participants. (Brunett, 1998) Given the geographical dispersions presented, *e.g.* from Maui to Virginia, special care needs to be taken to not add to the speed of light latencies, which are on the order of 120 milliseconds round trip.

Some of the exercises for which the U.S. warfighters require this capability would not be useful were they to not have 1,500 uniformed participants. These groups often included General Officers, so, any failure of the compute systems would present a significant loss time and an unacceptable cost to the government. This obviously calls for a fault tolerant architecture and very stable operations. Down time of even a few minutes would result in intolerable disruption and the loss of scarce time and financial resources. Further, the impact of even slightly increased latencies, $O(>200 \text{ milliseconds})$, on top of the speed of light latencies, would have the negative impact of making the simulations unacceptably unrealistic or even cause them to fail.

Meeting the needs of JFCOM for computational assistance had been focusing on procuring Linux Clusters via the U.S. Department of Defense High Performance Computing Modernization Program's Distributed Computing program and it was clear that HPCMP DC proposal requirements would occasion the continued evaluation of the need for and availability of additional capabilities.

While hardware and programming research was the major thrust of the research, part of the effort was driven by the need for designing, organizing, training and documenting the issues to produce the desired results. Research that was not made applicable by the users themselves would not be of use to them. In this case JFCOM did not have any resident HPC-trained personnel.

A continuing issue with distributed simulations has been the necessity for improved and effective logging processes. Logging was minimal and often after-action reviews were constrained to using human memory. Further, the distributed nature of the processing and analysis, made it difficult to effectively collate even the data that was available.

Another issue was the need for a capability for speed-up of analysis of the data. A delay in the analytic process is obviously unproductive *per se*, and the fact that many of the participants who could have benefited from the analytical data would be prevented from doing so by the requirement that they return to other duty immediately after the research exercises at JFCOM made the need for more rapid analysis necessary.

Part of the analytic problem was the perceived need to improve the Future After-Action Report System (FAARS) and its implementation. FAARS was a compute intensive activity and there seemed to be an opportunity to apply HPC systems to produce an improvement. This was made manifest by the requirements and requests of the users.

An issue that arose after the inception of the research project and was added to the mandates thereof was the need for accelerated HPC computing, which in this case took the form of a request for a 256 node, GPGPU-enhanced cluster. This need was exacerbated by the need to perform many calculations like Line Of Sight (LOS) and route planning. Without this acceleration, it was feared that the simulations would be unduly constrained.

As mentioned before, the systems were trans-continentially located, which, when coupled with the large amounts of data recovered, led to significant problems with transmission costs, delays in analysis, storage redundancies, and other data management issues. To address these problems, a new sub-effort was added to the research goals, a Distributed Data Grid or Scalable Data Grid.

Part and parcel of this same issue was the necessity to seamlessly and non-intrusively record important parameters during the research simulations and experiments. The need was to extract and make instantly available, data as to actions and impacts without succumbing too much to the Heisenberg Principal, *i.e.* to dramatically alter an object by the act of observing it. The resultant research effort was named the JLogger System. (Davis & Baer, 2006)

To sum up, the Joint Forces Command had and has an on-going need to utilize larger, higher resolution, more sophisticated, faster, and more fully populated urban simulations for training, analysis and evaluation. These simulations are live, virtual and constructive, as well as trans-continentially distributed. Size, speed, reliability and validity were and are constant concerns. (Davis, Lucas, Amburn & Gottschalk, 2005)

4. Methods

In general, the same method was applied to all of the above-enumerated problems. A careful analysis of current needs, existing systems, extant technologies and cost-benefit of varying solutions was undertaken. Working closely with the technical, administrative, instructional and analytical staff revealed opportunities for advancement and hurdles that could not be easily overcome. The research techniques applied were the time tested ones of applying a series of potential solutions and then assessing the resultant improvements as contrasted with the costs of implementation, in terms of hardware, software, maintenance, scalability and coding transparency.

Early work with JFCOM J9 user personnel had given the ISI researchers significant understanding of the capabilities of HITL simulation and JFCOM experiments. The major activity for the duration of the project was the continuous assessment of emerging requirements and developing technologies that might be germane in addressing the users' needs. Relying on methods developed by the ISI team members during their decades of experience with parallel processing and HPC system design, candidate technologies were identified, considered, implemented as deemed valuable, tested and evaluated for utility.

In approaching the need for a secure, effective nation-wide router network, the ISI team relied on early research they had conducted at ISI and Caltech since 1995 (Messina, 1998; Gottschalk, 1999). The current research revolved around how to most effectively implement that technology with existing data structures, how to provide needed enhancements to meet users' needs, and ascertain what those efforts meant to the community at large.

Identifying, validating and procuring the most appropriate systems were seen as the most productive way to ensure fault tolerant architecture and operations. When choices were presented to the ISI team and to JFCOM, the first priority was to select that option that would provide needed compute power, but would also yield the fewest disruptions to the users. This research goal differed slightly from that of a project in which any failure to computing was a real threat to life (weapon systems) and one in which the failure would have no untoward consequences (gaming systems). (Gottschalk, 2005)

One of the tasks of a researcher is to participate in the community in such a way as to provide the best opportunity for advancing research interests of all concerned. In this case, the ISI team felt the best provisioner of compute resources was the HPCMP who at the time of the initiation of the requirement was operating a Distributed Center program that had provided two 256 node Linux clusters already. Shortly after the beginning of the period of performance for this effort, the HPCMP changed that program's name to the Dedicated High performance computer Project Investment (DHPI) with similar opportunities to satisfy JFCOM's compute requirements. (Davis, 2008)

One of the new areas in this effort was the more formal presentation of training and documentation to enable and encourage the user's ability to fully utilize HPC technology. The research into the provision of these capabilities took the form of conceptualizing, designing, organizing, scheduling, presenting and evaluating three courses in the use of GPGPU technology. This evolution was intended to explore the extent to which this training could be developed, the extent to which the users would be interested, the flow of the needed information to the users, the degree to which the users were satisfied with the training and the level of utilization after the training. (Davis, Lucas, & Wagenbreth, 2008)

Approaching the logging process was straightforward conceptually, but held within its execution several areas, some more organizational than technical, which prior experience had indicated were problematic. Basically, the major impediment in the past was getting sufficient interaction with the users in the face of their total immersion in the spiral development process being driven by the needs of the "next HITL experiment" at JFCOM. Aside from that, the ISI method was to ascertain logging needs, data structures, hurdles to data access and organization, critical compute constraints and geographical/temporal needs, then try to optimize the data available and minimize the impact of data logging on simulation performance. (Graebener, 2003; Davis, 2006)

It was clearly an adjunct of the above to improve the speed of the analysis capability, but in addition to logging, the possibility of doing automatic preprocessing of the data to enhance the users' understanding of the import of the data available was kept in mind. One of the first aspects to be attacked was the transcontinental distribution of the data, which resulted in the design of the Distributed or Scalable Data Grid (SDG) discussed below. (Yao, 2005; Lucas, 2006)

One function of the analysis that was originally deemed to be important was the Future After Action Reporting System (FAARS) and its optimal implementation. The method here was to assess the ongoing needs of the users and analyze the best way to serve those needs. Again, it was the goal of the ISI team to bring their skills in parallel programming to bear on hitherto serially processed data to improve speed, data management, automatic analysis and data file production.

The on-going evaluation of JFCOM compute needs was undertaken. The first result of this was the decision to procure a 256 node GPGPU-enhanced cluster, which became a named work element of its own and is discussed below. Further, frequent meetings with the users and beating their needs against the ISI team's advanced knowledge of HPC technology was intended to yield an identification of potential computer enhancement. In addition to GPGPU acceleration, the ISI team considered the "Cloud Computing" concept, in relation to JFCOM needs. (Lucas, 2007)

As mentioned above, the trans-continentially arrayed nature of the data, computing assets and users called for a new capability that became known as the Distributed or Scalable Data Grid. The approach here was to carefully analyze data use, data location, user needs, bandwidth constraints, and applicable technology. This enabled data mining to be incorporated and analysis to take place *in situ*. This quickly led to the approach that a new system was needed and the application of data management techniques was required. (Yao, 2006)

In answering the needs of the user for improved logging, the development of the JLogger System was undertaken in the standard way again, i.e. the needs and constraints were analyzed and available or emerging technology was evaluated for relevancy and implementation plausibility, which encompassed not only initial coding, but ease of maintenance, modification and enhancement by the programmers at JFCOM. (Yao, 2007)

After assessing the needs of the users, it became apparent that a 256 node, GPGPU-enhanced cluster would be useful. The method adopted here was the creation of a document proposing such a research tool to the HPCMP. The effort was conceived as research into the use of GPGPUs as compute accelerators for a range of constraining computational requirements and then evaluating the utility of the cluster in JFCOM experiments. Secondly, there was the research issue into the productivity of the effort, mainly the degree to which journeymen operators could be trained and the degree to which they would use it for JFCOM experiments.

5. Assumptions

In identifying the assumptions made in this research effort, it would be prudent for the reader to remember that battlefield simulation and the compute requirements surrounding it are much more akin to the social than the physical sciences. So the overall assumptions for this work may be much more on the order of “... is assumed that the demographic information portrayed was valid and accurate ...” as opposed to a physical scientist’s assumption that “... a meter is the distance travelled by light in free space in $1/299,792,458$ of a second ($0,00000000033356409519815204957557671447492$ second).” The basic assumptions upon which this research was founded are:

- There is no real limit to the needs of the user in terms of
 - Size
 - Resolution
 - Sophistication
- Faster execution provides for more timely, *ergo* valid, analysis
- Careful parallel programming will prevent altering experimental validity
- The JFCOM mission is an important and justifiable one

These overall assumptions find their ultimate expression in the team’s view of the analysis of the capabilities for HITL and JFCOM experiments. Here the most over-riding assumption was that no matter how large, fast, populated, sophisticated and high-resolution a simulation was, the user could profit from and will soon ask for improvements. In the behavioral sciences, no different from the physical sciences, larger fields of view, more numerous active agents, faster execution, less constraining edge effects and shorter turn-around times all produce better results.

When considering the secure, effective nation-wide router network issue, the major assumptions were that there would be varying levels of security required, ranging from export controlled through SCI (Sensitive Compartmented Information.) Further, it was assumed that the team, while being provisioned by the DoD with secure encrypted networks, would have to operate under the constraints of security and conform to bandwidth requirements imposed by encryption devices.

The major issue driving fault tolerant architecture and operation assurance was the need to disrupt the users as rarely as plausible within economic and operational constraints and to provide for rapid recovery in the case of any failure, *e.g.* the loss of the Maui High Performance Computing Center (MHPCC) cluster during a storm in Hawaii. The assumptions here rested mainly on the quality and alertness of the operations personnel provided by JFCOM.

In interfacing with the HPCMP and considering the needs of the JFCOM users, the initial assumption was the continuing availability of the HPCMP’s DC program. This assumption also included the continued adherence to previously established procedures and proclivities in the awards process. These assumptions were marginally altered when a new program, the HPCMP DHPI program was announced. The compute requirements and the approach did not change appreciably.

In looking at the training and documentation effort, there were a number of assumptions going in. These ranged from assuming a reasonable level of C programming experience in the operator staff personnel to assuming they would be interested in training if presented in an adequate manner.

A similar approach was taken in the examination of the logging processes. The review of existing needs and future plans was conducted with the analysis of current systems and available technologies. Using a well-developed plan for synthesizing the results it was assumed that a careful balance of the needs and constraints would produce the final capability sought. Assumed goals were:

- Low levels of interference with current operations
- Rapid analysis of data
- Early availability of data
- Staged retrieval of data, *i.e.* most needed data earliest
- Low burden on network communications, including Wide Area Networks (WANs)

The parameters for the optimization of analysis capability speed-up were more difficult to isolate. The users had only vague notions of their needs and desires. In some instances, the ISI team members were able to suggest new capabilities, not conceived, perhaps not even conceivable, by the users. Here the underlying assumption really was, “A thorough understanding of the users’ needs and current capabilities led to the most relevant advances.” More germane assumptions were of the nature of the high likelihood of a dynamic system design and the need to make any proposed system easily maintained and practically adaptable.

As the project wore on, the major driving assumption about the FAARS implementation was the fact that it had a decreasingly important priority for the JFCOM users, ergo a concomitantly decreasing amount of time commitment from the ISI team. Many of the logging and Data Management research topics would be easily and productively applicable to the FAARS issues, should the JFCOM have a reawakened interest therein.

Assumptions for the 256 node, GPGPU-enhanced cluster initiative were manifold. They can be summarized as follows:

- Agent based simulation is in need of acceleration
- Line of sight (LOS) calculations are particularly amenable to GPGPU enhancements
- Route Finding would benefit from GPGPUs (Tran, 2008)
- GPGPUs are easier to program than Sony Toshiba IBM (STI) Cells or Field Programmable Gate Arrays (FPGAs)
- Parallel programming for a large GPGPU-enhanced cluster would allow scaling

These parameters all played a part in the decision. (Davis & Baer, 2006)

One of the emerging concepts not delineated specifically in the original research agenda was ISI's response to the data issues as was centralized in the Scalable Data Grid sub-effort. Without specific goals or definitive constraints having been laid out, the assumptions by the team were largely the framework in which the research was conducted. Some of the more important assumptions were:

- Data would continue to be trans-continentially distributed
- Data sizes would grow
- Archiving needs would emerge
- Speed and ease of access would be important
- Security would be provided externally, *i.e.* communications would be encrypted and sites would be secure

Finally, the JLogger System was another emerging technology, designed to respond to the logging issues mentioned above. The assumptions upon which this effort rested were as follows:

- Only a new system could provide the needed capability
- Success was tied to the reduction of impact on existing codes and operations
- Maintainability and utility were the driving factors

6. Procedures

During this project, ISI provided general research support to the J9 exercises. Due to conference and travel commitments, some of the monitoring was done, “on the road” from remote sites, sometimes literally from conference exhibit hall floors. This required a significant amount of operations sophistication to enable monitoring the status and availability of the three Linux Clusters: Koa, Glenn and the GPGPU Cluster Joshua.

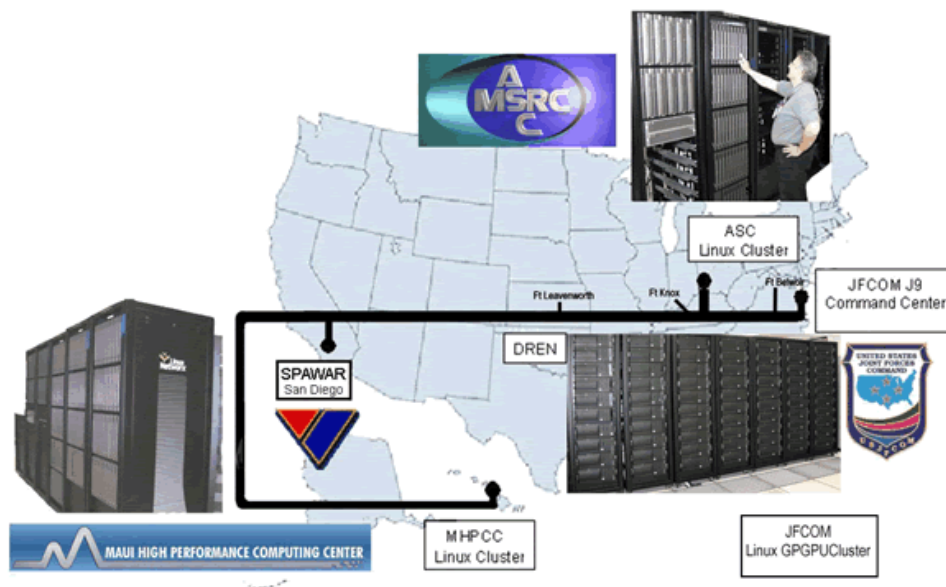


Figure 1: The JFCOM Clusters: Koa, Glenn and Joshua

Some staff were left in Marina del Rey, who could, should the situation have warranted it, made a last minute trip down to SPAWAR to gain classified access and Virtual Presence (ViPr) access.

As was earlier the case, when the ISI team participated in the operations and experiments themselves, there continued to be a considerable amount of effort that was expended in data management, logging, converting and storing as much as a Tera Byte every week during operations. The team also saw duty that required daytime monitoring and support for the data and evening transfers to the storage facility known as “Saber” in Suffolk, and then processing the data to store it in the standard format. To obviate the need of this bifurcated activity and to provide better real-time analysis, new concepts of operation were developed during this period for the SDG.

Again, the availability of the SPPs was one of the high points of each exercise/experiment. The “hot wash” was regularly scheduled, but in this period, the attendance by ISI personnel was almost invariably telephonic. As before, SPP reliability was outstanding. This reliability was a major thrust of the original work by the ISI team. (Lucas, 2003)

During this effort, substantial progress was also made in advancing toward the goal of providing an effective and secure, trans-continental router capability for JFCOM and their HPC operations. One of the first issues was a system test of the mesh-routers as an inter-node communications strategy and architecture. This was accomplished early in the period of performance.

Typical mesh-router tests were run on Koa, SPAWAR in San Diego, J9 in Suffolk and USC. The goals of the test were to identify and fix problems with the mesh-routers which would preclude their use in an event and to demonstrate to J9 personnel that the mesh-routers were stable and relatively problem-free. 128 unclassified nodes on Koa were used with a small number of puckers at J9 and SPAWAR. It was run as a federation with the tree routers and then run as the same federation with the mesh-routers. All personnel looked for any symptoms indicating that the mesh-router would not work as well as the tree router in an event. This was a prelude to creating network conditions and configurations in which the mesh-routers could deliver better scalability than the tree router.

Direction from JFCOM consistently put fault tolerant architecture and operations as a lower priority than other activities, so the only procedures implemented were to utilize native talent in both programming for stability and debugging to remove code problems.

In order to fully comprehend the needs of J9, ISI considered the experimentation to meet their needs in terms of magnitude and flexibility had previously been impossible due to limitations of compute power. An earlier DC award of the two clusters at Maui and Wright-Patterson AFB has enabled the development and implementation of a proven scalable code base capable of using thousands of nodes interactively. The JFCOM team continued to address community-wide issues such as: enhanced security for distributed autonomous processes, interactive HPC paradigms, use of advanced architectures, self-aware models, global terrain with high-resolution insets and physics-based phenomenology requisite for Joint Experimentation. The ISI team was instrumental in advancing this research agenda.

In wanting to address training, it was noted that the simulation community had often been hampered by constraints in computing: not enough resolution, not enough entities, and not enough behavioral variants. High Performance Computing (HPC) was held to be able to ameliorate those constraints. The use of Linux Clusters was advanced as one path to higher performance; the use of Graphics Processing Units (GPU) as accelerators was another. These are called General Purpose GPUs (GPGPUs). Merging the cluster and GPGPU paths was seen to hold even more promise. The ISI team members were the principal architects of a successful proposal to the High Performance Computing Modernization Program (HPCMP) for a new 512 CPU (1024 core), GPU-enhanced Linux Cluster, Joshua, for J9. This cluster was awarded to J9 via the Dedicated HPC Project Investment program, DHPI, and was configured in such a way to putatively utilize the GPUs to increase performance by a factor of two or more, with concomitant savings in cost, power and space.

Offering a course in GPGPU programming was seen as a way to aid JFCOM as they worked to take full advantage of the new JFCOM cluster. One of the unique aspects of the new machine was that there was a state-of-the-art NVIDIA GPU in each node. These GPGPUs could be programmed using the new CUDA programming language (Compute Unified Device Architecture, a "C-like" language). The DoD computing community in general needed to improve simulation performance and to make modifications to simulation programs such as JSAF, enabling them to take advantage of heterogeneous HPC architectures. Programming models, code examples and practice problems in CUDA were developed, drafted, documented and presented and test codes were implemented in class.



Figure 2: Course in GPGPU Programming - JFCOM, Oct 2007

Procedures for the improvement of the logging process entailed the application of careful study and analysis of the issues. The team checked-in a number of changes throughout the research effort for a set of sources that replicated and enhanced the presentation of the Contact Report. They thought that these classes represented the team's concept for how their overall system could be used. One of the major decisions was to use data cubes of various types. This entailed the creation of a number of new classes in a program subsequently called the Distributed or Scalable Data Grid.

Changes in the code caused continuing problems. The ISI team observed that in one of the events, the loggers were restarting every 60 seconds because the wrong parameters were given. So, at the beginning of each day it was decided to double check that the logger was running properly, and correctly archiving the binary data in gzip files. These corrective actions were successful and the logger continued to improve in stability and in utility. This raised the issue of the impact of the number of entities and how many could be run on a single

node of Koa or Glenn. In the observed events, the use of more than 120K - 200K culture entities resulted in simulations running out of memory, swapping (or expiring on one of the clusters). After one exercise, SPAWAR put down over 500k entities before performance became intolerable. The resulting memory use per simulation was about 2GB and CPU use was about 100% of a single CPU. The JFCOM operators thought 2 GB was the memory limit before swapping, even though there are 4 GB of memory.

The procedures applied for the analysis capability speed-up study were largely devoted to logging and to development of a tool for analysts to use in setting up scenarios. The use of GPGPUs was demonstrated on non-JSAF simulations, for ease of programming. The method employed was to use existing DOD simulation codes on advanced Linux clusters operated by JFCOM. The effort reported herein supplants the current JFCOM J9 DC clusters with a new cluster enhanced with 64-bit CPUs and nVidia 8800 graphics processing units (GPUs). Further, the authors have begun to modify a few legacy codes.

As noted above, the initial driver for the Forces Modeling and Simulation (FMS) use of accelerator-enhanced nodes was principally the faster processing of line-of-sight calculations. Envisioning other acceleration targets is easy: physics-based phenomenology, Computational Fluid Dynamics plume dispersion, computational atmospheric chemistry, data analysis, *etc.*

The first experiments were conducted on a smaller code set, to facilitate the programming and accelerate the experimentation. An arithmetic kernel from an MCAE “crash code” (Diniz, 2004) was used as vehicle for a basic “toy” problem. This early assessment of GPU acceleration focused on a subset of the large space of numerical algorithms, factoring large sparse symmetric indefinite matrices. Such problems often arise in Mechanical Computer Aided Engineering (MCAE) applications. It made use of the SGEMM (Single precision GEneral Matrix Multiply) algorithm (Whaley, 1998) from the BLAS (Basic Linear Algebra Subprograms) routines (Dongarra, 1993).

The GPU is a very attractive candidate as an accelerator for computational hurdles such as sparse matrix factorization. Previous generations of accelerators, such as those designed by Floating Point Systems (Charlesworth 1986) were for the relatively small market of scientific and engineering applications. Contrast this with GPUs that are designed to improve the end-user experience in mass-market arenas such as gaming. In order to get meaningful speed-up using the GPU, it was determined that the data transfer and interaction between the host and the GPU had to be reduced to an acceptable minimum.

Shortly after the beginning of the project, the technical personnel at JFCOM decided to de-emphasize the use of FAARS, so the ISI team did not expend an appreciable amount of time on the implementation of this system. JFCOM and AFRL management personnel concurred in this action.

One of the major benefits sought in this research was the deployment of a more effective programming model to support more effective simulation. The CUDA code from NVIDIA gives such an advantage because CUDA gives programmers the ability to exploit SPMD (single program multiple data) programming model on the GPU. CUDA programs are highly threaded. Access to shared memory space is achieved through gather and scatter operations. As per NVIDIA, here are two important factors concerning useful CUDA programs:

- There is no explicit synchronization mechanism with CUDA programs
- The wholesale executions in parallel are the real gain of GPU-based applications.

The literature represents that the compiler is a C pre-processor and specialized compiler that assists programmers with parallel programming. This would lead to the conclusion that computationally intensive sections should be explicitly tagged for execution on the GPU and are executed on segments of data on the GPU concurrently by many threads. Route finding is a class of algorithms that finds the “best” path, given a network of paths with N vertices (or nodes), between any number of vertices. The criteria for determining these paths (roads or edges) is determined by a cost function. The overall goal is to determine the min (or max) of the cost function of all of edges along the path.

Table 1 below summarizes some of the route finding algorithms. Note that the standard route finding algorithms apply to both the serial (CPU-based implementation) and parallel (GPU-based implementation). The ASSP algorithm is virtually the same as the SSSP algorithm, with the only difference being that it is implemented for all of the M paths in a particular network.

Table 1: Route-planning Algorithm Classifications

Algorithm Models	Properties			
	Implementation Model	Connectivity Graph	Storage Size	O(Time Complexity)
A* (A Star)	Serial	Priority Queue	N^2	$N \log N$
MM (Majorize-Minimize)	Serial & Parallel	Adjacency Matrix	N^2	$N^3 \log N$
FW (Floyd-Warshall)	Serial & Parallel	Adjacency Matrix	N^2	N^3
SSSP (Single Source Shortest Path)	Parallel	Adjacency List	N^2	$N \log N$
ASSP (Auction Sequential Shortest Path)	Sequen-Parallel	Adjacency List	$N^2 * M$	$N^2 \log N$

In Table 1 above, column A denotes the implementation model, B denotes the connectivity graph representation, C denotes the space (storage size), and D denotes the big O notation for time complexity.

In its utilized implementation, the JSAF simulator used the serial A* search algorithm to compute the “optimal” path for its clutter entities. The A* algorithm operated in $O(N \log N)$ time. The distributed federation-per-processor model implied that for every compute node, the JSAF load-balancer assigned a group of entities to its compute “basket.” The ISI design should restrict the computational space bound to these entities. Secondly, because JSAF is coarse-grain scalable, the ISI design exploits the higher resolution simulation per node by speeding up the computational time for the same amount of work on each node.

The A* algorithm is a heuristic implementation of the A algorithm. A* involves the use of heuristics to improve performance compared to the common perception of the greed of the BFS (breadth first search) algorithm, because the algorithm maintains a set of nodes (or vertices) already visited in a priority queue. The A* algorithm belongs to the class of single source single path family and it is by nature a serial algorithm, since only one node or vertex is considered at a time.

This procedural analysis was consistently followed in investigating new application areas of interest to JFCOM and amenable to GPGPU computation.

The development of the Distributed Data Grid was a real *de novo* project. Numerous data classes were developed. SimpleCube and SimpleCubeElement classes were established as a convenience to get something out of a Cube-Manager that could be used by "client" classes, in this case, the servlets. Creation of a CubeManager was a two-step process that created a factory for CubeDescriptions and then a separate step that created a factory for Cubes. The awkwardness of this approach was mitigated somewhat by having the servlets share a reference to a CubeManager. The new code was run on JESPP64, the twin AMD64 test machine that was housed at that time in the ISI JESPP Lab.

In addition, some convenience methods were added to the Cube interface and then it was found that they weren't needed. The required data was in the CubeView object with its collections of Dimension, DimensionNode, and DimensionValue objects. The test class CubeFactoryJdbcTest was also determined to be a good location to examine in order to see some low-level building choices. It was thought that all of the deprecated classes in the SDG OLAP directory were no longer used in anything other than "test" classes. It was decided to delete them from the Concurrent Version System (CVS) tree, after making adequate backup copies.

The need for a JLogger System was identified from the work above on logging needs by JFCOM. A design advanced by ISI was then adopted by JFCOM in the CVS as JLogger. JLOGGER was developed by the ISI team as a software system to support logging of “published” (as in “publish/subscribe”) data by JSAF and other federates that use Run Time Infrastructure (RTI) JLOGGER consisted of the following pieces:

- interceptor - library loaded at runtime by JSAF or other RTI federate to intercept published data and make it available to other JLOGGER software
- decoder - program which reads the interceptor output and saves it in various forms. It can send the data to a mysql which inserts it in a mysql database. It also writes it to disk. decoder can take input from these saved disk files instead of from the interceptor to facilitate offline after-action processing.

- aggregator - program which takes a single user input, sends it to multiple instances of mysql, reads and collates the results from the multiple mysqls, and sends the collated results back to the user. The user may be a program rather than a human user. Useful in distributed simulations.
- jloggerd - a collection of shell scripts to initiate, monitor, control and terminate the programs used in logging

A typical experiment might include sites participating from across the country. An example would be one in which the TEC site at Fort Belvoir, Virginia, had 30+ workstations and Saber, a quad-CPU machine with four TeraBytes of disk space that were used for after-event storage. The SPAWAR site at San Diego, California, had 20+ workstations. The J9 Joint Futures Lab at Suffolk, Virginia, had 50+ workstations and a 16-node mini-cluster. The ASC Wright Patterson Air Force Base at Dayton, Ohio, had the Glenn cluster with 128 dual CPU nodes. The MHPCC site at Maui, Hawaii, had the Koa cluster with 128 dual CPU nodes. This work was conducted prior to the acceptance of Joshua, the GPGPU-enhanced cluster.

These experiments typically ran five days a week, ten hours a day. Simulators might run all night, but with little activity and usually with logging disabled. Depending on availability and requirements, one or both of Glenn and Koa were used. Up to two hundred thousand clutter entities were simulated on the large clusters. (In this simulation, civilian entities are termed clutter, in that they serve to mask military entities.) Several thousand non-clutter entities were simulated on the other sites. A single node on the large clusters simulated 1000-2000 clutter entities.

7. Results and Discussion

While great strides were made in the number of entities to be simulated, the research implications of this achievement, albeit expected, were a dramatic reaffirmation of the value of HPCs. Nevertheless, on 14 December 2007, the Joint Forces Command Personnel, under the leadership of Rich Williams simulated a full ten million CultureSAF entities on the Baghdad terrain database. This was accomplished as is visualized in Figure 3 below:

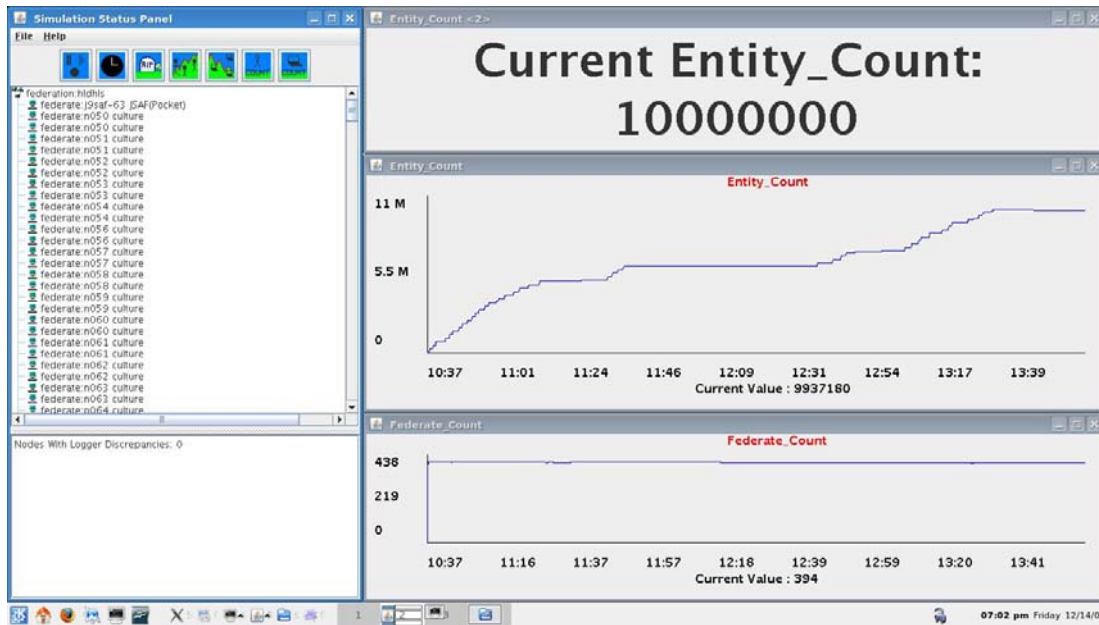


Figure 3: Graph of Entity Count Grow by Time on 14 Dec 07

This achievement had meaning in terms of the proposal that had been prepared by ISI and submitted to HPCMP by JFCOM, in which a major goal was the accomplishment of at least two million entities. A few early frustrations occurred due to the machine's not being configured for some of the experiments desired to help characterize the cluster's capabilities. Once again, the ten million entity run made the JFCOM, AFRL and ISI the largest agent-based, Semi-Automated Forces (SAF) implementation known, by at least two orders of magnitude, a size necessary to easily simulate any urban area with O(5M) inhabitants.

A summary of the results in the Mesh-Router tests showed that there were no problems with the mesh-router architecture or implementation. JFCOM operators and analysts confirmed that observed performance of the mesh-router was similar to performance of the tree router. Scalability was the desired parameter, so performance improvement was not expected nor required for success. There were no crashes or anomalies, a most critical factor.

The first effort was to create a baseline with the tree router which the mesh-router would match. The mesh-routers were incorporated in the latest RTI version. It typically took a long time to bring up a stable tree router configuration. Some of the problems the ISI team encountered were as follows:

- network problems
- clock skew between J9 and MHPCC inhibited kerberos authentication
- Multi-System Remote Control and Instrumentation (MARCI) problems with Runtime Initialization Data (RID) file and command line arguments that changed between RTI versions
- simulations that died immediately due to scenario input discrepancies
- cross-country network performance that was poor. This was solved by changing Transmission Control Protocol (TCP) to User Datagram Protocol (UDP).
- problem compiling loggers with new RTI. (an ongoing problem in this work was the use of upgraded compilers would fail until the code was re-written)
- applications were subscribing to too many entities, causing network problems.

JFCOM took down the tree routers on Koa. USC generated a new connectivity map and mesh map by hand. JFCOM replaced the tree routers on Koa with mesh-routers using the hand created maps. Map generation for mesh-routers needed to be implemented in MARCI. JFCOM put down 100K culture entities. System was stable and was left to run over several nights to test stability.

Upon checking in the morning, it was occasionally found that some of the tree router on Koa had died overnight for undetermined reason and had to be restarted. Koa mesh-routers and simulations remained up and stable all night. JFCOM added more clutter with the intent to stop at 500K or when the system died or response became poor. Problems with large objects and interest states had caused simulations to run out of memory quickly. These problems were not router related. Lack of swap space on Koa nodes made the simulation programs die rather than thrash.

Results for fault tolerant architecture and operations improvements were scant. During the period of performance, the need to conduct a test of Mesh Routers was raised. In order to better provide a trans-continental communications fabric for JFCOM use, the implementation of mesh routers, similar to the ones used on the local SPP meshes and the one that had been used in inter-computer communications from Maui to Aberdeen in the Synthetic Forces Express project (Brunette, 1998), was considered.

Before deploying mesh routers in an event/exercise, the team ascertained that there was a need to:

- 1) Merge mesh routers into Lockheed Martin CVS. The Caltech and the LMIS personnel were tasked to accomplish this.
- 2) Decide how they wanted to deploy the system. Two possibilities were surfaced:
 - a) Mesh on the SPP (ASC). Tree everywhere else.
 - b) Local trees on SPP, SPAWAR, TEC, J9. Mesh connecting the 4 sites.

Choice (a) was considered to be the least risky and demonstrated that the mesh routers work. There was no expectation of noticeable improvement in performance. Choice (a) required a minor upgrade to the mesh routers to command them, from the connectivity map, to initiate the connection to the tree router.

Choice (b) was determined to be the best use of the mesh routers. Depending on Defense Research and Engineering Network (DREN) routing capability, here there was some hope of seeing performance improvements. SPAWAR, TEC and ASC should be able to communicate without going thru J9.

- 3) It was decided to perform a test to establish confidence before use in an exercise. An unclassified test was suggested using some (30-60) Koa nodes, perhaps some ISI nodes, and whatever unclassified workstations are available and appropriate at TEC, SPAWAR and J9. Before that, it was recommended that a small shakedown test be performed, then LMIS or other personnel can participate to ensure confidence.

- 4) There are additions to the connectivity map format to specify mesh router capability.

- 5) Marci needed changes to the Graphical User Interface (GUI) to specify mesh router connectivity, and changes to generate the new connectivity map options.

Results for the careful consideration of the HPCMP DC compute requirements are clear, manifest and tangible: a 256 node cluster delivered in 2007.



Figure 4: Joshua

Success in the training and documentation area revolves around three courses taught at Suffolk, Marina del Rey and San Diego. In all more than 60 DoD personnel were trained in the use of a GPGPU-enhanced cluster. The fact that this capability is being put to use is reflected in the responses from the users, their contact with course instructors to resolve on-going issues and the attendance at technical papers presented discussing this work (Davis, 2007; Wagenbreth, 2007; Davis, 2009; Davis, 2009; and Lucas, 2009).

Results for the logging work were impressive. Data logging was performed in two modes, near real time and after action. Real time data was inserted in a SQLite database. (Graebener, 2003) A node simulating 1000 clutter items would generate a SQLite database of approximately 50 MB in an hour. The databases were deleted and reinitialized when they grew to over a gigabyte. If 100 nodes of the cluster were used for clutter simulators, approximately 5 gigabytes per hour of data was generated. For after action use, compressed binary data was stored in an archive directory. Binary compressed data is approximately $1/7^{\text{th}}$ the size of the corresponding database. Each night, the archived data was transferred to Saber, and expanded and decoded into a single MySQL database.

Clutter data from the Glenn and Koa clusters was not entered into the Saber database, due to size limitations. Data from 100 nodes on Glenn for a ten-day event would have been close to a TeraByte. Data from TEC, SPAWAR, J9 and J9 mini-cluster for non-clutter entities were entered into the MySQL database. Urban Resolve Phase I exercise generated about a TeraByte of data in the MySQL database.

The nightly data transfer was about 15 gigabytes of compressed data. Network transfer rate to Saber was approximately ten megabits per second. Three or four hours was required to do the transfer. Decoding and indexing the data into the MySQL database took 12 hours if everything worked perfectly. Human error and other factors usually prevented a day's data from being entered into the database before the next day's event started. It was usually at least several days after an event before the complete after action database was ready on Saber.

The logging routines used for the four exercises in the old configuration were adequate. It was the first attempt at logging data from hundreds of processors distributed geographically around the country simulating thousands of non-clutter entities. SDG is intended to remove deficiencies in the old methodology and upgrade what was essentially an experimental system into a production system. The design parameters for SDG specifically address the following list of deficiencies in the old system:

1. Near real time and after action data logging are implemented differently. Near real time queries are restricted by the use of simple aggregators.
2. The use of a single database on Saber does not have the capacity to include clutter data from the Glenn and Koa clusters.
3. Data transfers, decoding and indexing are time consuming and error prone, delaying the availability of the database. A goal is to have the complete database kept up to data continuously.
4. Retrieval of data and database generation for multiple exercises is inconvenient.
5. Expansion to more compute nodes, more entities per compute node and more data per entity is impossible. Disk storage, compute power, and network bandwidth all impose serious limitations.
6. The system does not respond gracefully to hardware and network problems. Saber is a single point of failure that makes all data unavailable.
7. Complex queries that may be useful to analysts are slow or impossible.

Database queries used in Urban Resolve are generally summary in nature. They count how many events or entities (database rows) meet specified criteria. Complex join operations were rarely, or never, used. Were it not for this constraint on the queries, an efficient distributed design would be more difficult.

A tool for the analyst to aid in setting up experiments was developed by ISI. Most of the work in this area was done in ISI's work on the code for a sensor route planning tool: Sensor PLanning and SCHEDuling tool (SPLASCH). It was designed to communicate with simulation federates, e.g. SLAMEM, which is an intelligence sensor simulation tool produced by Toyon. That product currently uses a TCP backdoor and Extensible Markup Language (XML). It will eventually use the High Level Architecture, (HLA) Run Time Infrastructure (RTI) to 'decode' JSAP data formats. This will produce a 'Sandbox' for algorithm development and demonstration of the Graphical User Interface (GUI.)

Characteristics include:

- Multiple platforms and sensors and depots
- Area Of Interest (AOI) types
- Points
- Circles
- Polygons
- Small *vs.* large – conceptually different
- Time constraints

It is clear that this would not be able to optimally solve general problems. There was an incipient need to determine ‘typical’ subsets on which to concentrate. The applied approach was to approximate the solution with a fast constructive algorithm, and then use iterative refinement with a ‘greedy’ uphill walk. A block diagram is included here as Figure 5.

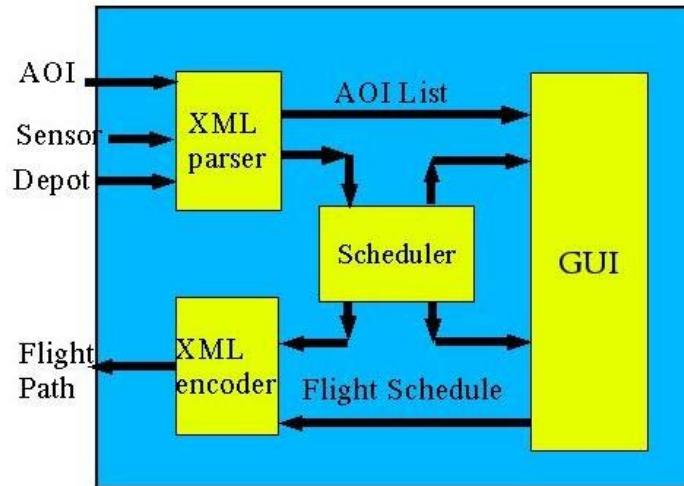


Figure 5: SPLASCH Block Diagram

It was reported that the planning tool required the following.

Input of the collection deck

- sample decks in XML format
- parsing the collection deck
- storing internal data structures

Ultimately, it was felt that knowledge of the following was needed:

- How to get the list of available platforms and sensors
- What are the allowed platforms and sensors, and their characteristics
- What are commonly used to know where to direct attention when designing and testing algorithms
- What are typical number and maximums for number of AOIs, number of sensors, number of sensor platforms
- If an AOI can be a point or a polygon
- Are there forbidden zones for the sensor platforms, due terrain or enemy counter measures

ISI created a graphical GUI to use as a 'sandbox' in designing and testing algorithms. Then current algorithms were menu/hand driven to see what worked. ISI was using a combination of multiple traveling salesmen, KMeans clustering and iterative refinement approaches.

ISI concluded that good progress was made in setting up a framework to develop and test. ISI required more specific information about AOI, sensor and platform characteristics. A GUI was tested to show sensor route planning at a notional level. It is represented here in Figure 6.

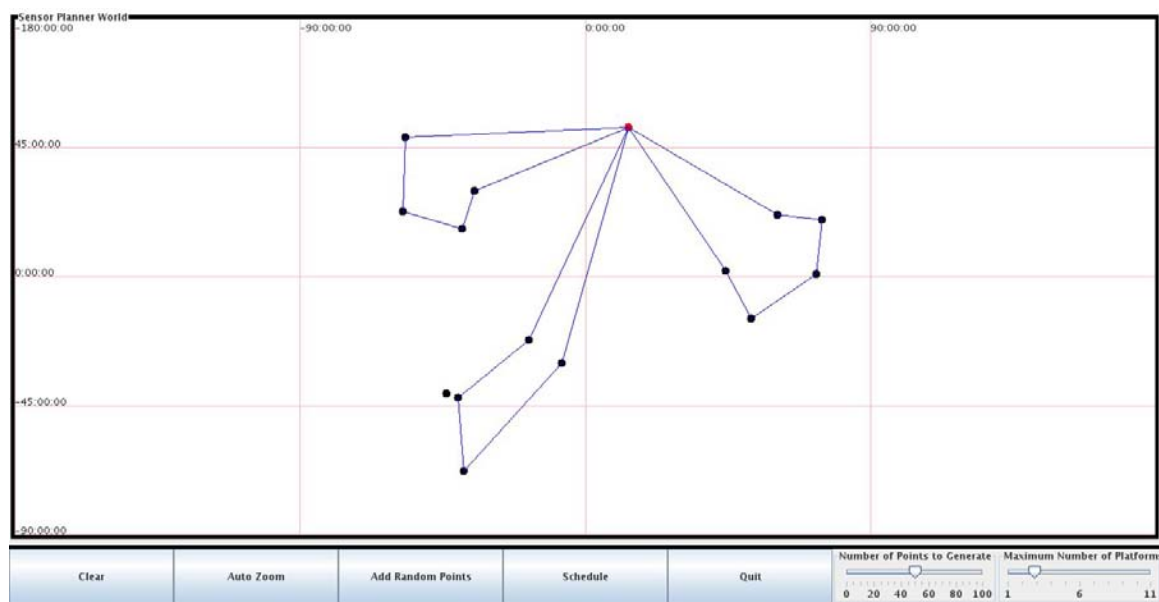


Figure 6: SPLASCH test GUI

The FAARS implementation had no formal results, as ISI and AFRL were directed by JFCOM to make this a low priority issue and repeated queries about its being researched were met with direction that research efforts were required elsewhere. Part of the conceptualization of this issue led to a study of Aggregation-De-aggregation (Gottschalk & Davis, 2006), but this line of research was not pursued further.

The results from the 256 node, GPGPU-enhanced cluster were very good. The initial year of research on the DHPI cluster Joshua was marked with typical issues of stability, O/S modifications, optimization and experience. All of the major stated goals of the DHPI proposal were met or exceeded. Research use by JFCOM was at a low level of operation due to issues outside the prevue of this re-port, but Joshua easily met its stability and availability requirements from JFCOM.

This particular project reflects the special needs of JFCOM. Instead of assessing the number of node-hours expended, the critical factor is the availability of the asset when required. Major General Larry Budge has commented that a nationally important intelligence asset has been fielded earlier and with significantly less cost due to joint experimentation using HPCMP assets.

Early work centered on the issues of getting the machine up and running. One problematic issue was get-ting the correct OS installed and coordinating that with the nVidia staff's recommendations as to varying version incompatibilities. This required careful coordination with the JFCOM SysAdmin personnel, who were invariably cooperative and professional. Characterization runs really began in earnest around December of 2007. The machine continued to be used as a development tool at that time.

The point-by-point report on the year's activities is as follows:

Joshua did provide 24x7x365 enhanced, distributed and scalable compute resources that did enable joint warfighters at JFCOM as well as its U.S. Military Service and International partners to develop, explore, test, and validate 21st century battlespace concepts in JFCOM J9's Joint Futures Laboratory (JFL). The specific goal is to enhance global-scale, computer-generated military experimentation by sustaining more than 10,000,000 entities on appropriate terrain with valid phenomenology. In addition, JFCOM J7 was also capable of prototyping global-scale, interactive supercomputer operation as part of JFCOM's new Joint Advanced Training and Tactics Laboratory (JATTL). Joshua was necessary to support the real-time, interactive requirements of the JFL and the JATTL.

The JFCOM team deployed existing DOD simulation codes which were previously run on advanced Linux clusters located at an appropriate site or sites, *e.g.* MHPCC and ASC-MSRC. The team supplemented the current JFCOM J9 DC assets with new clusters enhanced with 64-bit CPUs and graphics processing units (GPUs) in the form of Joshua. They began the process of modifying the legacy code to enable efficacious use of the new capabilities. As an important step in this procedure, the ISI team taught three GPGPU programming courses.

The results for the project to improve data management became known as the Distributed Data Grid project. (Yao *et. al.*, 2006) As part of the Scalable Data Grid (SDG) toolkit, ISI developed a prototype implementation of this agile data analysis framework. This framework was tested within the Urban Resolve 2015 exercises.

The team relied on a Meta-level Analysis Data Schema. In designing this implementation, the team strived to maintain flexibility. A central design element was a relational schema that implements the Analysis Data Model.

For efficiency reasons, data warehouse representation of multidimensional cubes typically use what is known as the star schema (Kimball *et. al.*, 1998). The star schema explicitly defines one relational table for each dimension. For the sensor/target scoreboard, the star schema would use two dimension tables to define the sensor dimension and the target dimension. One interpretation of the star schema is that it hard codes the two dimensions into the relational schema. Instead of hard coding, our approach is to define a meta-level schema that is capable defining and expressing multiple dimensions.

In this formulation, the sensor and target dimensions are defined as data, *i.e.* rows in the meta-level relational table. The `sdg_cube` table represents multiple dimensional cube definitions. Each cube is defined by an ordered list of dimensions. Each dimension has a name and an English description. Each dimension is defined by a set for concrete and abstract coordinates. These hierarchical coordinates form a partial ordering. Similar types of coordinates are grouped together and given a name. This is all described in greater detail in a paper presented at I/ITSEC, (Yao, *et. al.*, 2006)

The partial ordering of the coordinates also induces a partial ordering of the nodes. In a similar fashion we do not hard code measures, measure aggregation operators, and facts into fixed tables. Meta-level tables are defined to store these data models as data. The advantage of using a meta-model is that an analyst can easily and quickly design analysis data models tailored to his needs.

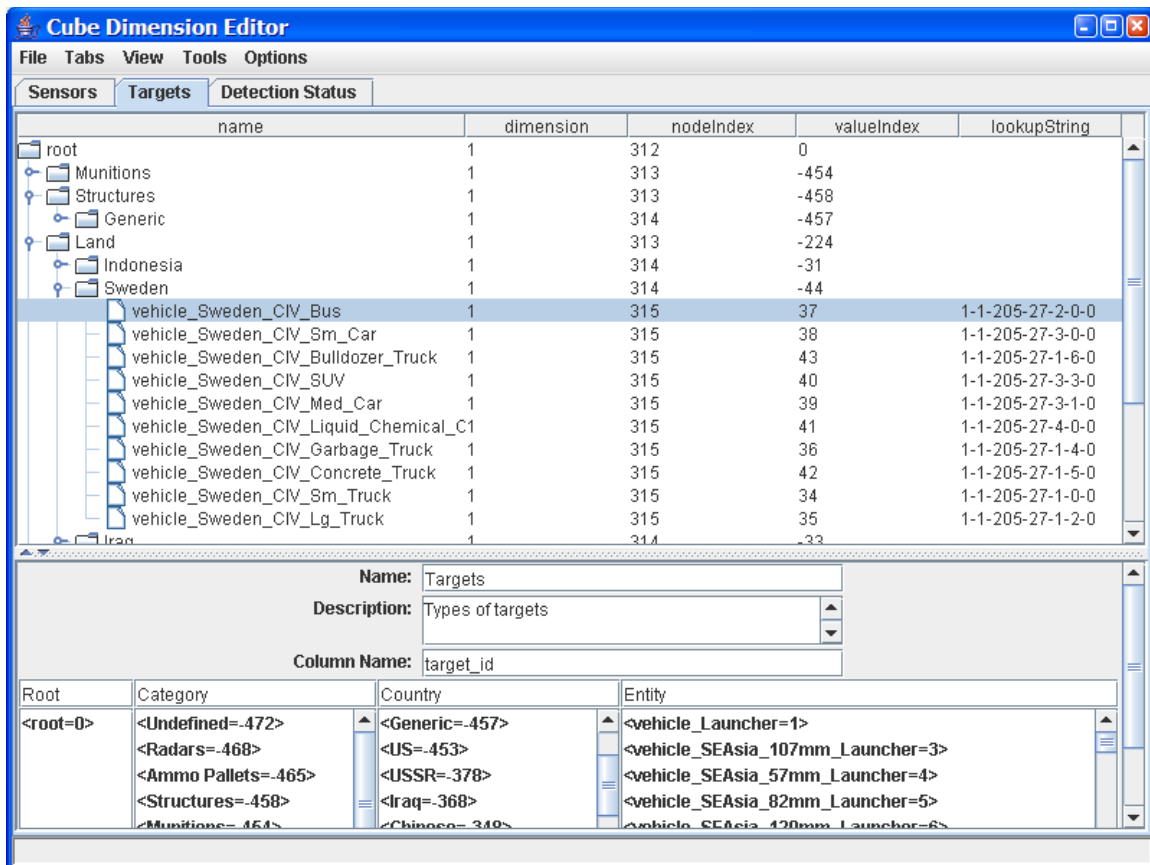


Figure 7: SDG's Cube Editor

Figure 7 shows a screen dump of the prototype SDG cube dimension editor. The editor presents to the user a tabbed view of the dimensions. Each tab corresponds to one dimension. The figure depicts a three dimensional sensor/target/detection status scoreboard. The detection status dimension breaks sensor contact reports down into four types: detected, not detected due to line of sight, not detected due to velocity, and not detected due to concealment. The top-half of each tab depicts the dimension as a tree-table. Each node/row in the tree-table represents a dimension coordinate.

For example, the entity vehicle_Sweden_CIV_Bus has coordinate 37, and the entity vehicle_Sweden_CIV_sm_car has coordinate 38. Both of these concrete coordinates belong to the abstract coordinate -44, called Sweden. Sweden in turn belongs to abstract coordinate -224, called Land. The tree-table has editing features that allows users to quickly define new coordinates or modify existing partial orderings. Editing operations include adding a new node/row; editing the content of a table cell; promote a node up the hierarchy; and demoting a node. Cut and paste operations are also defined. The editor uses Java JDBC to directly connect to relational databases in order to load and to store the dimension definitions.

The JLogger System continued during the contract and continues to this day to provide good service. It is currently being upgraded for the next round of experiments. In real time, an RTI interceptor captures published data and writes it to a Unix pipe. One decoder option is to read data from the Unix pipe and save it.

There are three ways the data can be saved:

1. Binary file
2. csv file
3. mysql database

If a user saves the data as a binary file during the event, he can use another decoder option after-action to read in the saved data and put it in a mysql database. There are numerous of options/command line arguments for the `jloggerd.sh` script, providing needed flexibility.

8. Conclusions

This research was able to identify and parameterize the capabilities for HITL and JFCOM experiments. This led to improvements. The early experiments on the University of Southern California Linux cluster (now more than 2,000 processors) showed that the code was scalable beyond 1,000,000 entities, given the availability of enough nodes (Wagenbreth, 2005). This effort is needed in order to deliver a tool set to military experimenters that they can use to easily initiate, control, modify, and comprehend any size battlefield experiment. It now additionally allows for the easy identification, collection, and analysis of data from these experiments, thanks to the work of Dr. Ke-Thia Yao and his team (Yao, 2005). The inherent scalability engendered in the JFCOM design of the computational system will allow sufficient computing power to be applied to each of these areas as needed.

The team's Mesh-Router technology was useful, stable, effective and non-disruptive. It did deliver a secure (via GFE Encryption), effective nation-wide router network. It was scalable, significantly more robust and showed great promise for future expansion. Further research is indicated and desirable. Once the initial configuration errors had been corrected, operation of the routers was indistinguishable from the previous communication network J9 had deployed. The Mesh-Router will enable J9 to maximize its use of network bandwidth while simultaneously reducing communication latency among the geographically distributed participants. Both the Mesh-Router and J9's earlier "tree" network are now available for J9 to use at its discretion.

One of the areas that did not receive much attention due to JFCOM priorities was the field of fault tolerant architecture and operations. It is our conclusion that some progress was made, but more remains desirable and possible. The strategy for creating a fault tolerant network will likely involve both dynamic reconfiguration as well as shadow routers. As discussed above, each has its place. The initial challenge will involve extending RTI-s to recognize directors and shadow routers. In addition, when a client has to reconnect following the loss of its router, it will also have to declare its entire interest space. MARCI will then have to be enhanced to incorporate them. Once these changes are made, J9 will have a practical level of fault tolerance in its communication network. The "fire and forget" message passing paradigm used within RTI means there will be packet losses when routers fail, but the number should be small enough that an experiment can continue.

The HPCMP DC proposal process and JFCOM compute requirements analysis was effective, as is represented by the award of a new cluster and the acceptance for publication of more than thirty papers directly tied to the use of the HPCMP assets. JFCOM's continued support of ISI research, again supports the conclusion that ISI effectively understood and responded to the JFCOM requirements *vis-à-vis* HPCMP.

It should be concluded that the ISI training was effective, based on the evaluation forms returned by the trainees. Those responses gave the ISI course some of the highest ratings ever seen by HPCMP. This leads the team to conclude that the training and documentation effort was both useful and germane to the users' needs.

The continued use of the logging processes developed indicates that the analysis of the processes were accurate and useful in designing new approaches and software. Only by living with the users in their environment could the ISI Computational Scientist rigorously examine, deeply understand and optimally respond to the users' needs.

Conclusion for the enhancement of analysis capability speed-up is more oblique and the proof thereof is more circumstantial. It could be argued that providing the tools for analysis capability speed-up was successful and opened up entirely new areas for future research. The SPLASCH system was a direct result of these efforts and the team's ability to respond to user needs with applied research and practical results.

It must be concluded that much was developed that would aid in FAARS implementation, but none of this was pursued, in accordance with JFCOM and AFRL direction. It remains to be implemented, should the need re-arise.

Conclusions regarding the 256 node, GPGPU-enhanced cluster, are that the JFCOM DHPI GPGPU-enhanced Joshua Cluster is a paradigm exemplar of leveraging technology to accomplish goals for orders of magnitude less funding. It can also be asserted that it enabled the analysis of systems in social environments that could not be disrupted by live exercises, e.g. downtown areas in U.S. urban environments. By emulating forces that would cost tens of millions of dollars to equip and deploy and by simulating urban areas that are not open to U.S. DoD exercises, JFCOM can now realistically, economically, safely and securely test new sensors, systems and strategies. HPCMP has achieved a preeminent position of professional leadership in the field of GPGPU-computing, showing the technical merit of the project. The computational merit of the project is clearly demonstrated in the achievement of the provision of adequate compute products to support several on-going exercises, one of which will be briefed. The authors carefully studied two algorithms, Line of Sight and Route Finding. Stability of Joshua in an operational setting will be explicated to show current progress. Appropriateness of requested resources in this case were a dead-on match, as Joshua has exceeded the goal of two million SAF entities by achieving ten million. The ISI team concludes that the GPGPU approach provided capabilities that could reduce purchase costs, enable large city simulations, save energy costs and deliver simulations reliably for the users.

Distributed or Scalable Data Grid conclusions are as follows: The ability to capture and log detail message traffic from very large scale simulations is now supported by our ability to analyze and comprehend that data. This work enabled a framework for quickly translating these "operational-level log data" into "analyst-level data." These data are capable of supporting decision makers. The framework explicitly defined a two-level data model that separates the operational logging data model from the analysis data model. The agility of the framework results from being able to isolate changes to the logging data model as a result of changes to the federation object model, and from being able to quickly define analysis data model that match analyst notion of measure of effectiveness and of performance.

The JLogger System is effective and continues to be used and appreciated by JFCOM users. Further work is ongoing, but the JLogger met its goals of logging, speed-up and non-interference with current operations. It has proved stable and easily maintained. Its continued use supports the conclusion that the analysis and the implementation were effective

The overall conclusion is that the JESPP0507 project met its goals, delivered more than expected, reacted well to changes in research direction and added considerably to the literature of this discipline.

9. Recommendations

An unending expansion of the needs and desires for increased capabilities for HITL and JFCOM experiments should drive future research in this area. All of the work done by ISI has been documented and published. The expertise generated is resident in the team members and is available to those who wish to benefit from it.

There will be a continuing need for a secure, effective nation-wide router network. In all of the paper presentations that were given, the issues most often raised by the audience are those of fault tolerant architectures and effective data management. JFCOM will continue to pursue these goals, which is, *a priori*, in accordance with any recommendation ISI has made on this issue. The reliance on GFE secure networks should not be taken to imply that there is not much yet to be accomplished in security. A goal should be to have secure communications over otherwise insecure media.

One of the future needs that will remain critical to DoD battle field simulations is that of fault tolerant architecture and operations. It is recommended that this area be highlighted, especially in light of the needs for high-bandwidth communications. Other DoD organizations are interested in this issue and, while JFCOM did not pursue this line of research in this effort, other agencies have funded ISI to do so and initial results have been achieved.

It is recommended that the analysis of JFCOM's HPCMP DC compute requirements continue, under the new program name DHPI. An operational need was identified for the Gulf Combat areas. This project was engaged in actively assessing the needs of both the Joint Urban Operations (JUO) and the Counter Mortar and Rocket Radar (CMR) problems. Each operations day presented new obstacles and new opportunities for high performance computing. Many of the simulation and data issues mentioned above speak to these challenges and opportunities. Direct and continuing collaboration with HPCMP is recommended.

One issue that flowed from the CMR problem was the lack of physics-based models that would be part and parcel of the JFCOM computational simulation suite. While look-up table, random number generated Monte Carlo simulations and existing code could emulate the battlefield to face validity standards useful for training, analysis called for more precise measures of what was occurring. A series of meetings was conducted with HPCMP personnel, Dr. David Pratt and Dr. Douglass Post, re the possibility of HPCMP's supporting research in the inclusion Forces Modeling and Simulation (FMS) into the new generation for computer and associated software of High Productivity Computing Systems (HPCS). Those discussions are on-going and should be pursued.

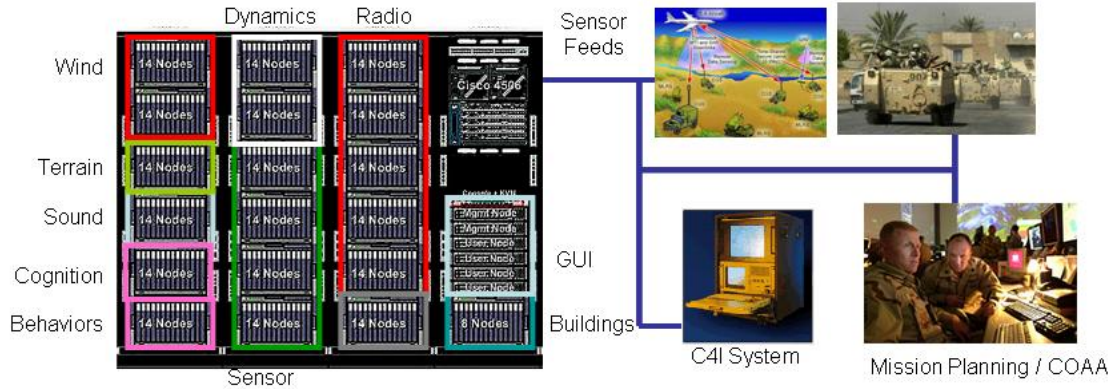


Figure 8: HPCS Computing Functionality Contributing to FMS

Future training and documentation awaits the definitions that only time will bring. At this juncture there are no plans for additional training courses, although some are being offered locally. The team remains professionally available to colleagues and DoD personnel for consultation without charge. It is recommended that the instructional capability be retained.

Recommendation for improving the logging process concern specific programming goals. Three very specific issues were isolated by the ISI programmers for future resolution:

- First, the team recommends that the systems and sub-systems be organized so there is a directory `rtis-1.3/meshrouter` that has all the files, including `meshrouter.cc` that has the main program. To look like the rest of the system this could be reorganized to make a `mesh-router` directory in the `lib` directory; and the source directory could be renamed `libmesh` or something similar. The main program could go in the `JSAP` tree in a new directory. Or, it could use the same main program used now, and `RTI` could check the connectivity map at runtime and see which kind of router is appropriate on a particular node.
- Second, it was hoped to find out how to make the code “telnet'able” like the tree router. It was also desirable to make all the commands work seamlessly with the rest of the system.
- Third, the team recommends further consideration of the uses of Message Passing Interface (MPI), in its multiple incarnations. There are multiple MPI implementations in the `mesh-router` code. It was the team's recommendation to convert all of the MPI instantiations to the `MPI-2` code.

Continuing to apply GPGPU acceleration is recommended as the most obviously productive area in the pursuit of analysis capability speed-up. The ISI team cannot but believe many outstanding advances are possible (Davis, 2009) in power and space savings. Additional emphasis on incorporating GPGPU code in the JSAF code base is recommended. Another area of assistance to the analysts was the SPLASCH program that has not been fully implemented. The analysts still claim they need this tool, so it is recommended that additional tasking be obtained to advance the work.

The team recommends that technology be implemented for use in the FAARS code at any time the users deem it useful. The dearth of research time allotted to this area precludes any other recommendations.

There can be no question about the use by JFCOM's J9 of the 256 node, GPGPU-enhanced cluster and its benefits to that organization. If the team had any recommendations it would be an expansion of the use of that machine and the J9 simulations to better allow the DoD to conduct training, analysis and evaluation.

The team is most adamant about recommendations for pursuit of the Scalable or Distributed Data Grid. In every public presentation, this was the item that attracted the most interest. Everyone in the DoD seems to be faced with a data glut, much of it as distributed as was JFCOM's. As concerned citizens, the team recommends some DoD organization be identified who will sponsor further research into this area.

Recommendations for the JLogger System are being carried out as this is being written. The improvements of the system and its continuous upgrade to allow it to be used on new simulations and during new experiments is vital to the interests of the DoD.

As far as general recommendations as to future research, the team cannot but believe that this is fertile ground indeed for continuing research, needing only HPC and Information Science research to bear important scientific fruit. The early results here confirm the validity of the approach and the benefits of the research. In terms of priority the team would suggest the Scalable Data Grid, GPGPU acceleration and the SPLASCH tools as those from which the most immediate and significant impact could be derived.

10. Management and Personnel Assigned

Personnel on the contract were as follows:

Dr. Robert F. Lucas, Dan Davis, Dr. Ke-Thia Yao, Gene Wagenbreth, John Tran, and Craig Ward are all from the Information Sciences Institute. Dr. Thomas is from Gottschalk Caltech.

11. References

- Brunett, S., Davis, D., Gottschalk, T., & Messina, P. (1998) Implementing Distributed Synthetic Forces Simulations in Metacomputing Environments, *The Seventh Heterogeneous Computing Workshop*, Orlando, FL.
- Charlesworth, A., & Gustafson, J., (1986), Introducing Replicated VLSI to Supercomputing: the FPS-164/MAX Scientific Computer, in *IEEE Computer*, 19:3, pp 10-23, March 1986
- Davis, D. & Lucas, R., (2006), Joint Experimentation, Data Management and Analysis Enabled by Trans-Continentially Distributed Linux Clusters, HPCMP Users' Group Conference, Denver, Colorado. June, 2006
- Davis, D. & Baer, G. (2006), "Providing Scalable Simulations, Environments and Data Structures for Theatre of War Analysis and Training," International Training and Education Conference, London, UK
- Davis, D. (2006), "Large-Scale Simulations and Transcontinental Data Management and Analysis", 4th Cluster Symposium On the Use of Commodity Clusters for Large-Scale Scientific Applications 2006, Baltimore, Maryland, July 2006
- Davis, D. & Moore, J., (2007), "FPGA and GPUs as Accelerators for Linux Clusters," in the Proceedings of the International Test and Evaluation Association M&S Conference, Las Cruces, New Mexico, 2007
- Davis, D., & Baer, G., (2005), "High Performance Computing Facilities for Joint Military Simulation Data Management," ITEA Modeling and Simulation Conference, Las Cruces, NM, 2005
- Davis, D., Gottschalk, T. & Davis, L., (2007), "High Performance Computing Allows Simulation to Transform Education," in the Proceedings of the Winter Simulation Conference, Washington, DC, 2007
- Davis, D., Lucas, R., Amburn, P. & Gottschalk, T., (2005), "Joint Experimentation on Scalable Parallel Processors," The ITEA Journal of Test and Evaluation, Fairfax, Virginia, 2005
- Davis, D.M., Lucas, R.F., & Wagenbreth, (2008), "Effectively Using a Large GPGPU-Enhanced Linux Cluster," in the Proceedings of the High Performance Computing Modernization Program Users' Group Conference, Seattle, Washington, 2008
- Davis, D.M., Lucas, R.F., Gottschalk, T.D., Wagenbreth, G., & Agaloff, J., (2009), "FLOPS per Watt: Heterogeneous-Computing's Approach to DoD Imperatives," in the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida, 2009

- Davis, D.M., Lucas, R.F., Wagenbreth, G., & Tran, J.J., (2008), "Effectively using a Large GPGPU-Enhanced Linux Cluster" in the Proceedings of the High Performance Computing Modernization Program Users' Group Conference, Seattle, WA, 2008
- Gottschalk, T. & Davis, D. (2006), "Application of Proven Parallel Programming Algorithmic Design to Aggregation-Deaggregation Problems," in the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida, 2006
- Gottschalk, T., Amburn, P. & Davis, D., (2005), "Advanced Message Routing for Scalable Distributed Simulations," The Journal of Defense Modeling and Simulation, San Diego, California.
- Graebener, R., Rafuse, G., Miller, R. & Yao, K-T., (2003), "The Road to Successful Joint Experimentation Starts at the Data Collection Trail". Proceedings of the Interservice/Industry Training, Simulation and Education Conference, Orlando, FL.
- Kimbal, R., Reeves, L., Ross, M. & Thornwaite, W. (1998) The Data Warehouse Lifecycle Toolkit. Hoboken, New Jersey: Wiley
- Lucas, R. & Davis, D., (2003), "Joint Experimentation in Scalable Parallel Processors," Proceedings of the Interservice/Industry Training, Simulation and Education Conference, Orlando, FL.
- Lucas, R., Davis, D. & Yao, K-T., (2006), "High Performance Computing Facilities for Joint Military Simulation Data Management," The ITEA Journal of Test and Evaluation, Fairfax, Virginia.
- Lucas, R.F., Davis, D.M. & Wagenbreth, G., (2007), "Implementing a GPU-Enhanced Cluster for Large-Scale Simulations," in the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida, 2007
- Lucas, R.F., Davis, D.M., Wagenbreth, G., & Tran, J.J., (2009), "Operational Use of a Large GPGPU-Enhanced Linux Cluster" in the Proceedings of the High Performance Computing Modernization Program Users' Group Conference, San Diego, California, 2009
- Tran, J.J., Lucas, R.F., Yao, K-T, Wagenbreth, G., Davis, D.M., & Bakeman, D.J., (2008), "A High Performance Route-Planning Technique for Dense Urban Simulations", in the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida, 2008
- Wagenbreth, G., Yao, K-T., Davis, D., Lucas, R., and Gottschalk, T., (2005), "Enabling 1,000,000-Entity Simulations on Distributed Linux Clusters," WSC05-The Winter Simulation Conference, Orlando, Florida.

- Whaley, R.C. & Dongarra, J.J., (1998), Automatically Tuned Linear Algebra Software, IEEE/ACM Conference on Supercomputing, SC98., pp.:38 – 38
- Yao, K-T. & Wagenbreth, G., (2005), “Simulation Data Grid: Joint Experimentation Data Management and Analysis,” in the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida, 2005
- Yao, K-T., Wagenbreth, G., & Ward, C.E., (2006), “Agile Data Logging for Large-Scale Simulations,” in the Proceedings of the Interservice/Industry Simulation, Training and Education Conference, Orlando, Florida, 2006

12. List of Acronyms

A*	A Star - A search algorithm by Hart, Nilsson and Raphael
AFRL	Air Force Research Laboratory
API	Application Programming Interface
ASC-MSRC	Aeronautical Systems Center – Major Shared Resource Center
ASSP	Auction Sequential Shortest Path
BFS	Breadth First Search
BLAS	Basic Linear Algebra Subprograms
CMR	Counter Mortar and Rocket radar
COTS	Commercial Off The Shelf
CUDA	Compute Unified Device Architecture
DAI	Data Access and Integration
DARPA	Defense Advanced Research Projects Agency
DC	Distributed Center
DCEE	Distributed Continuous Experimentation Environment
DHPI	Dedicated High performance computing Project Investment
DoD	Department of Defense
DQP	Distributed Query Process
DREN	Defense Research and Engineering Network
FAARS	Future After Action Report System
FPGA	Field Programmable Gate Array
Ft.	Fort
GB	Giga Byte
GFY	Government Fiscal Year
GigE	GigaByte per second Ethernet
GPGPU	General Purpose Graphics Processing Unit
GUI	Graphical User Interface
HITL	Human In The Loop
HLA	High Level Architecture
HPCMP	High Performance Computing Modernization Program
HPCS	High Productivity Computing Systems
ISR	Intelligence Surveillance and Reconnaissance
IAW	In Accordance With
JAWP HITL	Joint Advanced Warfighter Program Human in the Loop
J9	Joint Experimentation Directorate, JFCOM
JESPP	Joint Experimentation on Scalable Parallel Processors
JFCOM	U.S. Joint Forces Command
JSAF	Joint Semi Automated Forces
JUO-HITL	Joint Urban Operations-Human in the Loop
LMIS	Lockheed Martin Information Systems
MARCI	Multi-System Remote Control and Instrumentation
MCAE	Mechanical Computer Aided Engineering
MDS	Monitoring and Discovery System
MHPCC	Maui High Performance Computing Center
MM	Majorize - Minimize search algorithm
MOA	Memorandum of Agreement

MOLAP	Multi-dimensional On-Line Analytical Processing
OLAP	On-Line Analytical Processing
RID	Runtime Initialization Data
ROLAP	Relational On-Line Analytical Processing
RTI	Run Time Infrastructure
SAO	Situation Awareness Object
SCI	Sensitive Compartmented Information
SDG	Simulation Data Grid or Scalable Data Grid
SGEMM	Single Precision General Matrix Multiply
SIMD	Single Instruction Multiple Data
SLAMEM	Simulation of the Locations and Attack of Mobile Enemy Missiles
SPLASCH	Sensor Planning and Scheduling tool
SPMD	Single Program Multiple Data
SPP	Scalable Parallel Processors
SSSP	Single Source Shortest Path search algorithm
SQL	Structured Query Language
S/T	Sensor Target
STI	Sony, Toshiba and IBM (Consortium for Cell Processor game CPUs)
TDB	Terrain DataBase
U.S.	United States
ViPr	Virtual Presence (Video Conferencing System)
WAN	Wide Area Network

Appendix A: Presentations

Approved for public release; distribution unlimited

List of Presentations given by ISI on behalf of

**JESPP: Joint Experimentation
on Scalable Parallel Processors – 05-07
3rd Qtr GFY 05 thru 4th Qtr GFY 08**

Oral Presentations made during this effort

Paper Title (Journal Articles not Included)	Conference	Year
Effectively using a Large GPGPU-Enhanced Linux Cluster	HPCMP UGC	2009
FLOPS per Watt: Heterogeneous-Computing's Approach to DoD Imperatives	I/ITSEC	2009
Data Analysis for Massively Distributed Simulations	I/ITSEC	2009
Multi-Abstraction Levels in HPC: Enabling Consistency, Integration and Validation	HPCMP UGC	2009
GPU-Enhanced Linear Solver Results	SIAM PPSC	2008
A High Performance Route-Planning Technique for Dense Urban Simulations	I/ITSEC	2008
Effectively using a Large GPGPU-Enhanced Linux Cluster	HPCMP UGC	2008
Multi-Abstraction Levels in HPC: Enabling Consistency, Integration and Validation	HPCMP UGC	2008
A GPU-Enhanced Cluster for Accelerated FMS	HPCMP UGC	2007
Implementing New Educational Technology for a 21st Century DoD	I/ITSEC	2007
Modeling Human Performance of Situation Awareness in Constructive Experimentation	I/ITSEC	2007
Implementing a GPU-Enhanced Cluster for Large-Scale Simulations	I/ITSEC	2007
High Performance Computing Enables Simulations to Transform Education	WinterSim	2007
Incorporating High Energy Physics Data Capabilities into Large-Scale DoD Simulations	I/ITSEC	2007

High Performance Computing Enables Simulations to Transform Education	WinterSim	2007
FPGA and GPU Accelerators for Linux Clusters	ITEA	2007
Joint Experimentation, Data Management and Analysis Enabled by Trans-Continentially Distributed Linux Clusters	HPCMP UGC	2006
Agile Data Logging and Analysis for Large-Scale Simulations	I/ITSEC	2006
Successful Deployment of Mesh Routers: Trans-Continental JSAF Simulations	I/ITSEC	2006
Application of Proven Parallel Programming Algorithmic Design to Aggregation-Deaggregation Problems	I/ITSEC	2006
Educational Extensions of Large-Scale Simulations Enabled by High Performance Computing	I/ITSEC	2006
Interactive High Performance Computing: Lessons Learned	Cluster Symp	2006
Large-Scale Simulations and Transcontinental Data Management and Analysis	Cluster Symp	2006
Data Mining: New Insights from Distributed Data	ITEA	2006
Interactive High Performance Computing: Whence and Whither	ITEA	2006
Large-Scale Simulation Experimentation and Analysis Database Programming using Java	ITEA	2006
Developing Situation Awareness Metrics in a Synthetic Battlespace Environment	I/ITSEC	2005
High Performance Computing Facilities for Joint Military Simulation Data Management	ITEA	2005
Extending The Mesh-router Framework for Distributed Simulations	I/ITSEC	2005
Enabling 1,000,000-Entity Simulations on Linux Clusters	WinterSim	2005
Simulation Data Grid: Joint Experimentation Data Management and Analysis	I/ITSEC	2005
Joint Experimentation on Scalable Parallel Processors	ITEA	2005
Interactive HPC for Forces Modeling and Simulation: HPCMP's Distributed Center for JFCOM's JESPP Project	HPCMP UGC	2005
Scoping, Specifying and Satisfying the Need for Large-Scale Distributed Data Management *	Cluster Symp	2005
PetaScale FMS - Brief to DDS&T re FMS on HPC	HPCMP UGC	2005

Appendix B:

Abstracts of Presented Papers

Approved for public release; distribution unlimited

Extending The Mesh Router Framework for Distributed Simulations

Thomas D. Gottschalk

Center for Advanced Computing Research, Caltech
Pasadena, California
tdg@cacr.caltech.edu

Philip Amburn

SAIC, PET FMS On-site
Wright Patterson AFB, Ohio
philip.amburn@wpafb.af.mil

ABSTRACT

The Mesh-router system provides a general framework for scalable, interest-limited communications among processors in large-scale distributed simulations, such as the SAF family. The architecture was initially developed and implemented within the specific context of the ModSAF application and has recently been implemented in the JSAF/JUO application, using standard RTI-s communications primitives. This work provides a more general analysis of the Mesh-router system, clarifying the application-specific requirements for use of the communications framework and presenting a number of communications performance studies (total message throughput) for a system of simple federates using RTI-s communications. The overall Mesh-router architecture is reviewed, emphasizing the application-independent overall structure and the modest additional work needed to adapt the framework to the specific case of RTI-s communications.

The RTI-s Mesh-router is then compared with a tree-based communications built from standard RTI-s routers, using pair-wise message exchanges among simple federates. It is shown that Mesh-router performance is compatible with tree performance for trivial (e.g., nearest-neighbor) communications, and, more importantly, the aggregate bandwidth supported by the Mesh-router is substantially higher for non-trivial communications patterns, as would be expected in any realistic simulation environment. The communications performance studies are presented versus a number of relevant variables, including message size, total number of participating federates, and nominal length of the communications path. Extensions of the basic mesh topology used within the performance study are noted, including both modifications to support fault tolerance and a simple Tree/Mesh hybrid that could be easily implemented within the context of ongoing JSAF/JUO operations. Finally, the extensions of the existing Mesh-router software needed to support the OneSAF/RTI-N application are discussed.

Operational Experience and Findings: Distributed Simulations, Data Management and Analysis

Gene Wagenbreth

**Information Sciences Institute, USC
Marina del Rey, California
genew@isi.edu**

Ke-Thia Yao

**Information Sciences Institute, USC
Marina del Rey, California
kyao@isi.edu**

ABSTRACT

J9, the Experimentation Directorate of USJFCOM, and the Joint Advanced Warfighting Project of the Institute for Defense Analyses are conducting simulation experiments using operator workstations and hundreds of distributed computer nodes on Linux Clusters as a High Performance Computing solution to simulating hundreds of thousands of Joint Semi Automated Forces (JSAF) entities. A typical two-week experiment generates several hundred gigabytes of logged data. The data is queried in the near real time, and for months after an event. The amount of logged data and the desired performance of database queries motivated the redesign of the logger system from a monolithic database to a distributed database. The design of the distributed database incorporates several advanced concepts. Use of the distributed database in several two-week experiments presented significant challenges. Procedures and practices were established to execute the global-scale simulation, effectively use and monitor the distributed HPC assets, reliably and efficiently process and store hundreds of gigabytes of data, and provide timely and efficient access to the data via complex queries by analysts. This report describes the operation of the distributed database and the results obtained. It further discusses the development of effective techniques to identify, diagnose and resolve various impediments to efficient operations. Data is presented to support the choices made and future work is discussed.

Simulation Data Grid: Joint Experimentation Data Management and Analysis

Ke-Thia Yao and Gene Wagenbreth

Information Sciences Institute
University of Southern California
Marina del Rey, California
kyao@isi.edu , genew@isi.edu

ABSTRACT

The need to present quantifiable results from simulations to support transformational findings is driving the creation of very large and geographically dispersed data collections. The Joint Experimentation Directorate (J9) of USJFCOM and the Joint Advanced Warfighting Project is conducting a series of Urban Resolve experiments to investigate concepts for applying future technologies to joint urban warfare. The recently concluded phase I of the experiment utilized and integrated multiple scalable parallel processors (SPP) sites distributed across the United States from two supercomputing centers at Maui and at Wright-Patterson to J9 at Norfolk, Virginia. This computational power is required to model futuristic sensor technology and the complexity of urban environments. For phase I the simulation generated more than two terabytes of raw data at rate of >10GB per hour. The size and distributed nature of this type of data collection pose significant challenges in developing the corresponding necessarily data-intensive applications that manage and analyze them.

Building on lessons learned in developing data management tools for Urban Resolve, we present our next generation data management and analysis tool, called Simulation Data Grid (SDG). The two design principles driving the design of SDG are 1) minimize network communication overhead (especially across SPPs) by storing data near the point of generation and only selectively propagating the data as needed, and 2) maximize the use of SPP computational resources by distributing analyses across SPP sites to reduce, filter and aggregate. Our key implementation principle is to leverage existing open standards and infrastructure from Grid Computing. We show how our services interface and build on top of Open Grid Services Architecture standard and existing toolkits, such as Globus. SDG services include distributed data query/analysis, data cataloging, and data gathering/slicing/distribution. We envision the SDG to be a general-purpose tool useful for a range of simulation domains.

Developing Situation Awareness Metrics in a Synthetic Battlespace Environment

Jacqueline M. Curiel, Michael D. Anhalt

Alion Science and Technology

Fairfax, Virginia

jcuriel@alionscience.com, manhalt@alionscience.com

John J. Tran, Ke-Thia Yao

Information Sciences Institute, USC

Marina del Rey, California

jtran@isi.edu, kyao@isi.edu

ABSTRACT

The Joint Forces Command (JFCOM) conducts Joint Urban Operation (JUO) exercises in synthetic battlespace using human-directed computer simulation tools such as Joint Semi-Automated Forces (JSAF) to support ongoing joint war-fighting efforts. A component of these experiments is that of human-in-the-loop (HITL) interactions where human players impact the outcome of the exercise. This is in contrast to Monte Carlo constructive experiments that only involve computer behavior. The need to objectively measure the effectiveness of human players and their interaction with the simulation environment requires quantitative metrics to supplement more qualitative observer-based judgments. Situation awareness (SA), a cognitive behavior captured in HITL experiments, involves the perception and comprehension of forces and events in a situation, and a prediction of their future status, Endsley (1995). Objectively measuring SA is drawing intense interest because this knowledge is crucial to successful decision-making processes (C2). Building upon work presented at I/ITSEC 2004 (An Interdisciplinary Approach to the Study of Battlefield Simulation Systems, paper 1886), we adopt a cognitive-computational approach for measuring SA based on situation model theory. Situation models are complex mental representation of events. As events unfold, these mental representations must be updated to maintain an accurate representation. Prior research has demonstrated that situation models are updated along a number of dimensions. These dimensions reflect information about entities, space and time coordinates, participants' goals, and the causal relationships of events. We utilize the information encapsulated in SA objects (SAOs), recorded during the JUO exercises, to develop a tool that automatically monitors players' SA and evaluate the importance of these dimensions on situation awareness over the time course of the experiment and on the three levels of SA. Our findings have practical implications for subsequent training, product development, and extend the knowledge base of cognitive behavior.

Incorporating High Energy Physics Data Capabilities into Joint Forces Simulations

Julian J. Bunn and Thomas D. Gottschalk
Center for Advanced Computing Research, Caltech
Pasadena, California
julian@cacr.caltech.edu, tdg@cacr.caltech.edu

ABSTRACT

The data management and data exploitation issues for large-scale, distributed DoD simulations have striking parallels within a number of existing large-scale High Energy Physics (HEP) projects, in particular, the experiments associated with the Large Hadron Collider (LHC) in Geneva, Switzerland, which will begin operating in 2007. The significant commonalities include: data rates of 10-100 GBytes/day, data distribution and database operations over very large scale, high-speed networks, and sophisticated data exploitation objectives. In this regard, the lessons learned over the past decade of preparations for LHC operations have obvious significance and relevance for operational (fielded) DoD information exploitation systems. These similarities are substantial, in spite of some significant differences between the DoD and HEP applications. In particular, the distributed data generation within typical DoD experiments (e.g., JUO or CMR) is quite unlike the massive single point of data generation within an LHC experiment.

This paper explores three particular areas of DoD data exploitation needs having significant parallels with existing HEP/LHC work. The first involves robust, scalable database design and management, such as the distributed simulation and data system within the Joint SemiAutomated Forces project now under development within the US Joint Forces Command. Important aspects here include operational transparency and efficiency from the perspective of a single user/analyst at a workstation. The second general area involves support for “user toolkits” - significant additional computational subsystems such as data-mining/knowledge-discovery procedures and “what if” Monte Carlo excursions that go well beyond straightforward queries of a distributed data base. The final area has to do with “real-time” considerations, where this term is to be understood in the more general sense of legitimate, possibly urgent user needs that exceed available computational resources. Strategies are discussed for leveraging the demonstrated HEP expertise toward DoD data management and exploitation problems.

Extending The Mesh-router Framework for Distributed Simulations

Thomas D. Gottschalk

Center for Advanced Computing Research, Caltech
Pasadena, California
tdg@cacr.caltech.edu

Philip Amburn

SAIC, PET FMS On-site
Wright Patterson AFB, Ohio
philip.amburn@wpafb.af.mil

ABSTRACT

The Mesh-router system provides a general framework for scalable, interest-limited communications among processors in large-scale distributed simulations, such as the SAF family. The architecture was initially developed and implemented within the specific context of the ModSAF application and has recently been implemented in the JSAF/JUO application, using standard RTI-s communications primitives. This work provides a more general analysis of the Mesh-router system, clarifying the application-specific requirements for use of the communications framework and presenting a number of communications performance studies (total message throughput) for a system of simple federates using RTI-s communications. The overall Mesh-router architecture is reviewed, emphasizing the application-independent overall structure and the modest additional work needed to adapt the framework to the specific case of RTI-s communications.

The RTI-s Mesh-router is then compared with a tree-based communications built from standard RTI-s routers, using pair-wise message exchanges among simple federates. It is shown that Mesh-router performance is compatible with tree performance for trivial (e.g., nearest-neighbor) communications, and, more importantly, the aggregate bandwidth supported by the Mesh-router is substantially higher for non-trivial communications patterns, as would be expected in any realistic simulation environment. The communications performance studies are presented versus a number of relevant variables, including message size, total number of participating federates, and nominal length of the communications path. Extensions of the basic mesh topology used within the performance study are noted, including both modifications to support fault tolerance and a simple Tree/Mesh hybrid that could be easily implemented within the context of ongoing JSAF/JUO operations. Finally, the extensions of the existing Mesh-router software needed to support the OneSAF/RTI-N application are discussed.

High Performance Computing Facilities for Joint Military Simulation Data Management

Dan M. Davis

Information Sciences Institute, USC
Marina del Rey, California
ddavis@isi.edu

Garth D. Baer

George Washington University
Washington, District of Columbia
garthbaer@gmail.com

ABSTRACT

A widespread problem is the overwhelming amount of output data inundating many in the simulation user community. Much of this torrent is generated by current high-end computational capabilities. Especially joint and combined forces analysts are faced with the two major tasks of first validating and then utilizing the data generated by modern techniques. A major part of the solution is an optimized software architecture. To enable that effort to achieve success commensurate with the users' goals, a dedicated and appropriately designed data management facility was required. Taking cognizance of the advances made in the physical sciences' community, such a facility was conceived, designed and is being proposed to the HPCMP. The techniques of identifying, quantifying and incorporating important data handling parameters required for success should be applicable to many large data set problems.

This paper will discuss the general state-of-the-art in data management, the specific problems presented by the U.S. Joint Forces Command simulations of up to a million independent SAF entities on a global-scale terrain, the methods used defining the problems presented thereby, and the path to the decision to stand up a new facility. Using the successful techniques found effective in originally generating the information, e.g. studying approaches used by other scientific research efforts, effective data management schemes have been discovered. Both the design process and the architecture itself will be laid out. Some issues addressed will be the choice of compute platform, the provision of associated communications, the selection of storage peripherals, the analysis of incipient technical advances that are likely to be germane, cost-benefit analyses of competing installations and the approach necessary in order to design for the future. Specific performance, cost and operational issues will be presented and analyzed. Lessons learned from this evolution should be extensible into many fields associated with modeling and simulation.

Advanced Message Routing for Scalable Distributed Simulations

Thomas D. Gottschalk

Center of Advanced Computing Research,
California Institute of Technology
Pasadena, CA 91125
tdg@cacr.caltech.edu

Philip Amburn

SAIC, PET FMS On-Site
Wright-Patterson AFB OH 45433
philip.amburn@wpafb.af.mil

Dan M. Davis

Information Sciences Institute, USC
Marina del Rey, CA 90292
ddavis@isi.edu

On large Linux clusters, scalability is the ability of the program to utilize additional processors in a way that provides a near-linear increase in computational capacity for each node employed. Without scalability, the cluster may cease to be useful after adding a very small number of nodes. The Joint Forces Command (JFCOM) Experimentation Directorate (J9) has recently been engaged in Joint Urban Operations (JUO) experiments and Counter Mortar analyses. These both required scalable codes to simulate over a million SAF clutter entities, using hundreds of CPUs. The JSAF application suite, utilizing the redesigned RTI-s communications system, provides the ability to run distributed simulations with sites located across the United States, from Norfolk, Virginia to Maui, Hawaii. Interest-aware routers are essential for scalable communications in the large, distributed environments, and the RTI-s framework, currently in use by JFCOM, provides such routers connected in a basic tree topology. This approach is successful for small to medium sized simulations, but faces a number of constraining limitations precluding very large simulations.

To resolve these issues, the work described herein utilizes a new software router infrastructure to accommodate more sophisticated, general topologies, including both the existing tree framework and a new generalization of the fully connected mesh topologies. The latter were first used in the SF Express ModSAF simulations of 100K fully interacting vehicles. The new software router objects incorporate an augmented set of the scalable features of the SF Express design, while optionally using low-level RTI-s objects to perform actual site-to-site communications. The limitations of the original Mesh-router formalism have been eliminated, allowing fully dynamic operations. The mesh topology capabilities allow aggregate bandwidth and site-to-site latencies to match actual network performance. The heavy resource load at the root node can now be distributed across routers at the participating sites. Most significantly, realizable point-to-point bandwidths remain stable as the underlying problem size increases, sustaining scalability claims.

Keywords: Linux, cluster, scalable, JSAF, routers, Communications.

Joint Experimentation on Scalable Parallel Processors (JESPP)

Dan M. Davis, Robert F. Lucas

Information Sciences Institute, USC
Marina del Rey, California
ddavis or rflucas@isi.edu

Philip Amburn

Science Application Int. Co.
Wright Patterson AFB, Ohio
philip.amburn@wpafb.af.mil

Thomas D. Gottschalk

California Institute of Technology
Pasadena, California
tdg@cacr.caltech.edu

ABSTRACT

The JESPP project exemplifies the accessibility and the utility of High Performance Computing for large-scale simulations. In order to simulate future battlespaces, US Joint Forces Command's J9 required expansion of its JSAF code capabilities: number of entities, behavior complexity, terrain resolution, infrastructure features, environmental realism, and analytical potential. Synthetic forces have long run in parallel on networked computers. The JESPP strategy exploits the scalable parallel processors (SPPs) of the High Performance Computing Modernization Program (HPCMP). SPPs provide a large number of processors, interconnected with a high performance switch and a collective job management framework. JESPP developed software routers that replaced multicast with point-to-point transmission of interest-managed packets. This article lays out that design and development. It also details several events that have simulated up to one million clutter entities, which were "fought" from Suffolk, VA. These entities were executed on remote SPP's, one in Maui and one in Ohio. This paper sets forth the authors' experience in scoping the hardware needs, developing the project with HPCMP, and implementing the system.

Petascale Computing for Military Operations

David R. Pratt, Ph.D.

Philip Amburn, Ph.D.

Science Applications International Corporation (SAIC)

12901 Science Drive

Orlando, FL 32817

(407) 243-3308

prattda@saic.com

Robert F. Lucas, Ph.D.

Dan M. Davis

Information Sciences Institute

University of Southern California,

4676 Admiralty Drive

Suite 1001

Marina del Rey, CA 90292

Keywords:

Military Simulation, HPCS, Petascale

ABSTRACT: The traditional combat models we have employed to date can no longer represent current military operations. The reasons for this are threefold: limited scale, insufficient fidelity, and inadequate combat focus. Ironically, all three share a common root cause; namely, lack of processing ability. The legacy codes in use have had to make compromises in order to operate within the distributed processing environments for which they were developed. At the same time the models have been proliferating, military operations have become significantly more integrated, thus increasing the gap between simulations and operations. There is now a confluence of events that provide a dramatic opportunity for the use of new high productivity computer systems (HPCS) computing systems for the Department of Defense (DoD). Investments by US JFCOM, PEO STRI, DARPA and HPCMO are already coupling HPC resources with operational needs to support the radical transformation of the US military. HPCS-level resources can provide exciting new capabilities to the warfighter by combining HPC-based functional, physical, logical, and behavioral models of battlespace components and effects in a human-in-the-loop application. But it needs to be done in a disciplined manner. By taking advantage of the convergence of the processing capabilities of the HPCS resources, the component nature of emerging simulations, such as OneSAF Objective System (OOS), and the location transparency provided by the long haul networks, we propose to replace the selected component elements of OOS architecture with either high-fidelity, first order physics models or proxy interfaces to operational systems. In doing so, we are replacing the areas that traditionally have been most simplified by the computational and network limitations of the distributed processing model with those elements most needed to emulate current military operations.

While the reduced cost of determining the war fighting impacts of various resource allocations is a major benefit of Forces Modeling and Simulation (FMS) on a HPCS-level resource, the most significant benefit is closing the gap between simulations and operations. More realistic training, experimentation, analysis, and planning will lead to a reduction in casualties and an increase in mission effectiveness. With the complexity of the modern and future battle space, this can only be done on a HPCS class resource.

Application of Proven Parallel Programming Algorithmic Design to the Aggregation/De-aggregation Problem

Thomas D. Gottschalk

Center for Advanced Computing Research, Caltech
Pasadena, California
tdg@cacr.caltech.edu

Dan M. Davis

Information Sciences Institute, Univ. of So. Calif.
Marina del Rey, California
ddavis@isi.edu

ABSTRACT

A continuing problem in entity-level, intelligent agent simulations has been one of efficiently, effectively and expediently aggregating smaller units like squads and platoons into larger ones like companies and battalions and then de-aggregating them again at appropriate times. This paper reviews the goals and issues of the aggregation/de-aggregation (A/DA) problem and then lays out some solutions based on High Performance Computing, computational science and lessons learned from advanced techniques, such as adaptive simulation meshes. Experience has shown and logic dictates that aggregation is a more straightforward operation than is de-aggregation. A/DA of collective units is required for future, large-scale simulations, *e.g.* Sentient World Simulation. Understanding how to distribute the smaller units and how to represent the impacts of the simulation on these segments has largely eluded the M&S community for years. This problem is made more complex by the existence of significant amounts of “legacy code” and this paper gives examples of a successful approach to working with such code in an HPC environment. Three workable solutions are enabled by HPC: simulating all levels continuously while displaying only the designated unit level, simulating smaller entities’ behavior with reduced behavioral resolution to save compute resources, and foregoing all lower level simulation by simulating only the top-level designated. This last method requires laying down the lower-level entities using doctrine, status, and terrain to achieve realistic disposition. This paper will investigate the processes, impacts, and performance of all three methods. Entity migration across various compute nodes in cluster computers and germane HPC examples from similar computational approaches will be described. The approach applies methods, shown to be effective in on-going research in the physical sciences, to problems facing the DoD M&S community. Performance analyses are anticipated, as are user evaluations by operators, controllers, and analysts.

Line of Sight and Route Planning Performance Using Advanced Architectures

Gene Wagenbreth and Robert Lucas

Information Sciences Institute

Marina Del Rey, California

genew@isi.edu, rflucas@isi.edu

ABSTRACT

Current computers usually include a Graphics Processing Unit (GPU). The arithmetic processing capability of these GPUs generally exceeds the capability of the computer's central processing unit (CPU) by an order of magnitude or more. Use of the GPU as an arithmetic accelerator has been discussed by Dinesh Manocha, UNC, and others (IITSEC 2005). The GPU is difficult to program and the calculations to be performed must fit certain criteria in order to use the GPU effectively. This paper examines the feasibility of utilizing these results in the JSAF code in the Urban Resolve experiments. The Joint Semi Autonomous Forces (JSAF) simulation software is used to model hundreds of thousands of entities. Available processing power limits the number of entities simulated on a single CPU. To determine the value of a GPU for JSAF in urban terrain, we looked at two algorithms that utilize a significant portion of the processor capability. These are Line of Sight (LOS) and Route Planning calculations. Both algorithms are contained in a small portion of JSAF source code. This makes translation to GPU code possible. The LOS calculation, particularly when approximated, maps very well onto a GPU. The approximation is such that "can not see" calculations are exact. "Can see" calculations must be recalculated exactly on the base CPU. The Urban Resolve trials use terrain dominated by buildings and roads, in contrast to other experiments dominated by natural terrain. In order to determine the feasibility of moving LOS and Route Planning to the GPU, JSAF was instrumented to continuously measure the time spent on these tasks. "Can see" and "can not see" results from LOS were separately instrumented. This paper presents the results of running instrumented JSAF in scenarios commonly used by JSAF. A modified LOS approximation algorithm is presented which may allow more efficient execution.

Successful Deployment of Mesh Routers for Trans-Continental Urban Resolve Experiments

Thomas D. Gottschalk

CACR, Caltech
Pasadena, California
tdg@cacr.caltech.edu

Robert Lucas, Gene Wagenbreth, and John Tran

Information Sciences Inst. /Univ of So Calif
Marina del Rey, California
{rflucas, genew, jtran}@isi.edu

William Helfinstine

Lockheed Martin Info
Burlington, Mass
bill.helfinstine@lmco.co

ABSTRACT

This paper reports the successful deployment of a robust scalable interest-managed router architecture that has supported a series of trans-continental simulations, such as Urban Resolve. Previous architectures had served well over the years, but were conceptually limited both in scalability and in robustness, or fault-tolerance. The scalable router architecture had its inception in high performance parallel computing research and its initial application in a truly scalable architecture for inter-node communications on parallel supercomputers and Linux clusters. Its design provided both needed scalability and desirable robustness on the single platform meshes of several large parallel computers made up of hundreds of compute nodes. The scalable router was designed to integrate smoothly with other Urban Resolve software by reusing Run Time Infrastructures (RTI-s) components. In an effort to minimize communication latency, maximize use of available network bandwidth, and increase robustness of trans-continental (Virginia to Hawai'i) operations, Joint Forces Command's J9 directed that its wide-area router's offer the same characteristics of scalable and robust operations. That led to the wide-area deployment of the scalable routers. This paper sets forth the experience of that evolution, the non-disruptive incorporation of the new routers, the scalability of the interest-managed routing, and the performance of the new network. The assiduous factorization of the program, in order to optimize and temper the code, bore fruit during the implementation process and that factorization activity is explicated and analyzed. Further, the authors look to their experiences in high performance computing to lay out future capabilities and directions for additional development. The area of primary interest and importance is fault tolerance. A specific proposal for the design and fielding of a system impervious to the loss of individual router processes is presented.

Agile Data Logging and Analysis

Ke-Thia Yao, Gene Wagenbreth and Craig Ward

Information Sciences Institute
University of Southern California
4676 Admiralty Way, Marina del Rey, CA 90292
{kyao, genew & cward}@isi.edu

ABSTRACT

The High Level Architecture Object Model Template (HLA OMT) supports simulation interoperability by providing a Federation Object Model (FOM) to formally describe the information interchange (objects, object attributes, interactions, and interaction parameters) within a simulation federation. Information used by a single federate within the federation is defined by the Simulation Object Model (SOM).

Often the federate SOMs are mutually incompatible, so standing up a federation typically requires a tedious process modifying the simulation federates to conform to the proposed FOM. A variety of agile FOM techniques have been proposed to facilitate this integration process.

From the simulation data logging and analysis perspective, there is an analogous problem of adapting the analysis tools to particular federations. Data analysis tools are designed in accordance with the analysts' notion of Measures of Effectiveness (MOE) and Measures of Performance (MOP). Often these measures are not directly compatible with respect to the underlying federation object model. This is especially troublesome for the lower-level MOP, which must have common characteristics with the logged FOM data.

This paper presents a two-layered framework that supports the agile adaptation of analysis tools to specific federations. The top semantic layer provides a modeling framework to capture concepts that analysts tend to use. The concepts include measurements and dimensions. Examples of dimension include are object classifications, time, and geographic containment. The lower syntactic layer describes how to map the particular federation object models to more abstract semantic concepts. In addition, we show how this approach supports reuse by taking advantage of the hierarchical nature of the object models. These concepts are now being successfully implemented and evaluated in the Joint Forces Command Urban Resolve 2015 experiment.

Modeling Human Perception of Situation Awareness During Constructive Experimentation

John Tran
ISI
Marina del Rey, CA
jtran@isi.edu

Philip Colon, Brian Mainz
Toyon Research Corporation
Goleta, CA
{philipc,bmainz}@toyon.com

Jacqueline Curiel, Michael Anhalt
Alion Sciences and Technology
Fairfax, VA
{jcuriel,manhalt}@alionscience.com

ABSTRACT

Highly advanced sensor technologies give our military commanders a significant command and control (C2) advantage over our enemies during conflicts. Similarly, in a synthetic battlespace the use of advanced sensor technology models, such as the Simulation of the Location and Attack of Mobile Enemy Missiles (SLAMEM), gives human-in-the-loop (HITL) participants parallel advantages. There are two accepted simulation methodologies for analyzing the impact of sensor technologies: (1) through HITL experiments, such as Joint Urban Operations (JUO), and (2) through Monte Carlo constructive (MCC) simulations. For HITL experiments, which are dominated by human interaction and behavior, all three levels of situation awareness (SA), Endsley (1995), can be derived from situation awareness object (SAO) encapsulation. MCC experiments, which by design lack any human interaction, are dominated by algorithmically determined behaviors. Sensor measurements can be fused to perceive individual entities, but currently lack the capability to recognize groupings of entities. This behavior is a partial perception of the first level of SA. Furthermore, sensor data fusion models lack the capability to automatically recognize the second and third orders of SA, function and intent, respectively.

The paper will report on research into the development of synthetic SAOs (SSAO) that can be incorporated into MCC runs. These synthesized objects must be sufficiently expressive to capture the three levels of SA, and have an initial condition based on SA metrics collected from HITL experiments. Furthermore, the data attributes correlated with SSAO in the large-scale distributed experiments are statistically compared to ground-truth data collected from the JFCOM's JUO and Counter Mortar and Rocket (CMR) HITL experiments. Using this approach, erroneous assumptions the players made while creating SAOs will be replicated algorithmically. This leads to quantifying and better understanding player deviations (variance in human activity) during HITL experiments and improving human interaction when designing sensor models.

Interactive High Performance Computing: Lessons Learned

Dan M. Davis

ddavis@isi.edu

Information Sciences Institute
University of Southern California
Marina del Rey, California

From the beginnings of warfare, the leader has been challenged to optimize his force structure through analysis, intelligence, training and control. The early sand table has morphed into the sophisticated simulation on powerful HPC assets such as Linux Clusters. These simulations' usefulness is accepted; their benefits only slightly mitigated by difficult programming paradigms; their contributions acknowledged by top DoD leaders. All these, however, would come to naught were it not possible to establish and maintain a suitable interactive environment within the HPCMP. The HPCMPO is committed formally to making this a reality and this paper will speak to challenges, successes and future opportunities of advancement.

The problems can be laid out as falling within three broad concept areas:

- Interactive operations within a batch-oriented community
- Security assurance across a distributed-simulation net
- Computer security in a multi-user (100s) and machine-controlled environment.

Each of these has emerged as a real problem, with incipient "show-stopping" consequences. Not infrequently, these issues become critical with significant human assets are being kept idle, awaiting resolution. The obstacles are described in some detail, both to give the audience a flavor for what has happened, but also to provide a plausible plan for progress in similar situations.

The computer operations environment in which this transformation is occurring is the Distributed Center (DC) awarded to the U.S. Joint Forces Command (JFCOM) by the HPCMP in 2003 and delivered in April of 2004 to MHPCC and to ASC-MSRC. The DC consisted of two Linux Networx, 128 node clusters (dual 3.0 GHz Xeons, 2 GB memory, 60 GB HDD, GigE Internode comms.) The need calling for this new DC was the experimental requirement of JFCOM to populate its urban warfare environments with up to one million independent agent civilians, blue forces, red forces and associated environmental phenomenology. This experimental environment was used to assess the efficacy of newly conceived sensor systems, principally via the SLAMEMTM program. These experiments have garnered the attention of high-ranking officers (*e.g.*, Gen. Abizaid and ADM Giambastiani) and political leaders (*e.g.*, Senator Clinton and former Speaker Gingrich.)

But implementing these experiments ran counter to the usual practices of the established HPC community, which have been optimized over a decade of use for batch operations. Several unique and innovative solutions were sought and introduced to enable an interactive operational paradigm. Personnel at both centers and at JFCOM worked diligently to enable this vital capability. Yet, some issues remain unresolved and will likely require high-level intervention to advance the state of the art in interactive high performance computing and FMS computational science.

One of the early issues to need resolution was how to handle security when several trusted sites (GenSer Secret) were using a program that initiated and controlled processes on remote computers without Kerberized user intervention, after secure card log in by the senior system administrator at the simulation site. Another major issue was the provision of help-desk and SysAdmin services on a real-time basis when problems arose during operations when upwards of 150 operators were sitting at their terminals waiting and Major Generals were calling for action. This is in opposition to the typical batch operation, with which this community is so familiar, where responding in hours or days is considered adequate and genteel.

An example of an issue still ripe for resolution is the need for a close look at the need for traceability of individual programmers during the tumult of operations in a simulation bay, where many operators will need to log-in, use terminals for a few key-strokes, then move on, all the while requiring system privileges at the root level. While not compatible with the current operating modality of most HPC centers, it is requisite to achieve the goals of the JFCOM user.

Data Mining:

New Insights from Distributed Data

Ke-Thia Yao

Information Sciences Institute, Univ. of Southern California
4676 Admiralty Way, Suite 1001, Marina del Rey CA 90292
Phone: 310 448-8297 - FAX: 310 823-6714 kyao@isi.edu

Data mining is, by loose consensus, the extraction of useful new information from data that was designed to be collected for some other purpose. In the fields of Test and Evaluation (T&E) and of Forces Modeling and Simulation (FMS), incredible amounts of information are collected for very focused purposes. Within this data there lie many important insights, unperceived, but not imperceptible, the value of which remains to be apprehended. The author's experience in teaching an introductory course on Data Mining at the Viterbi School of Engineering at the University of Southern California has resulted in a new vision for the applicability and utility of Data Mining to T&E and FMS environments.

A quick overview of the theory, implementation and use of data mining will be given. Specific applications in T&E and FMS will be adduced to give practical examples. Future uses, probable products and visions of cohesive approaches will be discussed.

Joint Experimentation, Data Management and Analysis Enabled by Trans-Continentially Distributed Linux Clusters

Dan M. Davis and Robert F. Lucas

ddavis@isi.edu & rflucas@isi.edu

Information Sciences Institute
University of Southern California
Marina del Rey, California

ABSTRACT **Of Oral Presentation** **For JFCOM Distributed Center for FMS**

There is a well-defined and critical need to produce quantifiable results, which are derived from simulations, to support transformation experimentation of the Joint Forces Command (USJFCOM). This is driving the creation of very large and geographically dispersed data collections. The Joint Experimentation Directorate (J9) of USJFCOM and the Joint Advanced Warfighting Project are conducting a series of Urban Resolve experiments to investigate concepts for applying future technologies to joint urban warfare. The recently concluded phase I of the experiment required, utilized and integrated multiple scalable parallel processors (SPP) sites distributed across the United States. These were hosted by the supercomputing centers at Maui (NHPCC) and at Wright-Patterson (ASC-MSRC) on a net including J9 at Suffolk, Virginia, Topographic Engineering Center, Fort Belvoir, Virginia and SPAWAR San Diego, California. This computational power had to be harnessed by scalable code in order to model the capability of futuristic sensor technology and the complexity of the urban environment. For phase I the simulation generated more than two terabytes of raw data at rate of >10GB per hour. The size and distributed nature of this type of data posed significant challenges in developing the corresponding data-intensive applications that manage and analyze them. Building on lessons learned in developing data management tools for earlier Urban Resolves, a next generation data management and analysis tool, called Simulation Data Grid (SDG), was developed and implemented. The design principles driving the architecture of SDG were

1. minimize network communication overhead (especially across SPPs) by storing data near the point of generation and only selectively propagating the data as needed, and
2. maximize the use of SPP computational resources and storage by distributing analyses across SPP sites to reduce, filter and aggregate.

The key implementation principle was to leverage existing open standards and infrastructure from Grid Computing. The developed system services interface with and were built on top of Open Grid Services Architecture standard and existing toolkits (Globus). SDG services include distributed data query/analysis, data cataloging, and data gathering/slicing/distribution. It is argued that SDG has proven to be a general-purpose tool useful for a range of simulation domains.

Interactive High Performance Computing: Whence and Whither

Dan M. Davis

Information Sciences Institute, Univ. of Southern California
4676 Admiralty Way, Suite 1001, Marina del Rey CA 90292
Phone: 310 448-8434 - FAX: 310 823-6714 ddavis@isi.edu

Interactive High Performance Computing (HPC) is essential for the future of battlespace simulation. This and other fields will benefit from real-time use of HPC capabilities, but this opportunity will come to naught if it is not possible to establish and maintain a suitable interactive environment within the HPC Community. The defense of the Nation may very well depend, to some degree, on making this Interactive HPC a reality and this paper will speak to challenges, successes and future opportunities of advancement.

The problems can be laid out as falling within three broad concept areas:

- Enabling interactive operations within a batch-oriented community
- Assuring security across distributed-simulation environments
- Facilitating multi-user (100s) and machine-controlled secure computing

Each of these has emerged as a real problem, with nascent “show-stopping” consequences. Not infrequently, these issues become critical when significant human assets are being kept idle, awaiting resolution. The obstacles will be presented in some detail, both to give the audience a flavor for what has happened, but also to provide a plausible plan for progress in similar situations.

A computer operations environment in which this transformation is occurring is the Distributed Center (DC) awarded to the U.S. Joint Forces Command (JFCOM) by the HPCMP in 2003 and delivered in April of 2004 to MHPCC and to ASC-MSRC. The DC consisted of two Linux Networx, 128 node clusters (dual 3.0 GHz Xeons, 2 GB memory, 60 GB HDD, GigE Internode comms.) The need calling for this new DC was the experimental requirement of JFCOM to populate its urban warfare environments with up to one million independent agent civilians, blue forces, red forces and associated environmental phenomenology. This experimental environment was used to assess the efficacy of newly conceived sensor systems, principally via the SLAMEM™ program. These experiments have garnered the attention of high-ranking officers (*e.g.*, Gen. Abizaid and ADM Giambastiani) and political leaders (*e.g.*, Senator Clinton and former Speaker Gingrich.)

But implementing these experiments ran counter to the usual practices of the established HPC community. These practices have been optimized, over a decade of use, for batch operations. Several unique and innovative solutions were sought and introduced to enable interactive operational capabilities. Personnel at both centers and at JFCOM worked diligently to enable this vital capability. Yet, some issues remain unresolved and will likely require novel approaches to advance the state of the art in interactive high performance computing and Forces Modeling and Simulation (FMS) computational science.

One of the early issues to need resolution was how to handle security when several trusted sites (GenSer Secret) were using a program that initiated and controlled processes on remote computers without Kerberized user intervention, after secure card log in by the senior system administrator at the simulation site. Another major issue was the provision of help-desk and SysAdmin services on a real-time basis when problems arose during operations when upwards of 150 operators were sitting at their terminals waiting and Major Generals were calling for action. This is in opposition to the typical batch operation, with which this community is so familiar, where responding in hours or days is considered adequate and genteel.

An example of an issue still ripe for resolution is the need for a close look at the need for traceability of individual programmers during the tumult of operations in a simulation bay, where many operators will need to log-in, use terminals for a few key-strokes, then move on, all the while requiring system privileges at the root level. While not compatible with the current operating modality of most HPC centers, it seems requisite to achieve the goals of the JFCOM user.

Educational Extensions of Large-Scale Simulations Enabled by High Performance Computing

Dan M. Davis

Information Sciences Institute, USC
Marina del Rey, California
ddavis@isi.edu

Laurel K. Davis

Apple Tree Services
Culver City, California
laurel.davis@appletreeservices.com

Summer L. Allen

Science Applications Int'l. Corp.
Alexandria, Virginia
sallen@objectsciences.com

ABSTRACT

Large-scale intelligent agent simulations, enabled by high performance computing (HPC), have been effectively used by the Department of Defense for experimentation and analysis. The authors analyze their experiences in these and related areas, then present data and conclusions to support new applications of proven pedagogies to broaden the value of these capabilities across the areas of training and education. Over more than a decade, HPC has shown the ability to enable otherwise unattainable sizes of intelligent agent simulations, growing from small unit, to battlefield, to theater of war, and, finally, to global-scale operations. The techniques necessary to achieve these levels were imported and adapted from early supercomputing research in basic science projects at major universities. Among the insights from that research were the reductions of validity and utility suffered when constrained samples of the subject phenomena were simulated. This paper extends that concept into the discipline of education and demonstrates the putative desirability of having large-scale capabilities in the educational environment as well. The authors describe the available technologies for large-scale simulations, review the successes of experimentation and analysis enabled by those technologies, and outline the many opportunities for implementation in education. They then focus on early experimentation using distributed HPC to aid in technical and non-technical education for all age cohorts. They lay out a roadmap for future development and for assessments of applicability of their techniques by others who should benefit from such capabilities. Cost/benefit analyses are invoked to assist the potential users in making valid evaluations of the applicability of these proven techniques to their own uses. The development of an interactive educational module is outlined, described and lessons learned are reported. A test on a trans-continental meta-computing platform will be reported from the viewpoint of both HPC performance and educational efficacy.

Modeling Human Perception of Situation Awareness During Constructive Experimentation

Philip Colon

Toyon Research Corporation
Goleta, CA
philipc@toyon.com

John Tran, Ke-Thia Yao

Information Sciences Institute
Marina del Rey, CA
{jtran, kyao}@isi.edu

Michael Anhalt, Jacqueline M. Curiel

Alion Science and Technology
Marina del Rey, El Cajon, CA
{manhalt, jcuriel}@alionscience.com

ABSTRACT

Highly advanced sensor technologies give our military commanders a significant command and control (C2) advantage over our enemies during conflicts, particularly with respect to situation awareness (SA). The use of advanced sensor technology models in synthetic battlespace gives war fighters parallel advantages. Two accepted simulation methodologies for analyzing the impact of sensor technologies are through Human-in-the-Loop (HITL) experiments, such as Joint Urban Operations (JUO), which utilize sensor capabilities to assist human participants during the experiments, and Monte Carlo Constructive (MCC) simulations, which can be used to model human performance. In HITL experiments using Joint Semi-Automated Forces (JSAF), participants describe their SA using Situation Awareness Objects (SAOs) which then can be reconstructed using Endsley's (1995) three levels of SA (perception, comprehension, and prediction). MCC experiments, which are dominated by algorithmically determined behaviors, can be used to model SA. Sensor measurements currently can be fused to perceive individual entities, but do not have the capability to recognize groupings of entities, resulting only in partial perceptual SA. Furthermore, current sensor data fusion models do not produce the second and third levels of SA, comprehension and prediction. This paper will report research efforts to utilize both methodologies to expand the use of SAOs beyond player declarations to the automatic generation of SAOs. We develop a method to organize events drawn from scenarios taken from HITL experiments using SAOs in order to develop situation awareness algorithms for the MCC runs. These model-generated synthetic SAOs (SSAOs) can be compared to SAOs generated by human players to identify the accuracy of the models as well as be used to identify strengths and weaknesses in player performance.

Implementing New Educational Technology for 21st Century DoD Leadership Development

Laurel K. Davis

Next Generation Leaders, Inc.

Culver City, California

laurel.davis@appletreeservices.com

Dan M. Davis

Information Sciences Institute, USC

Marina del Rey, California

ddavis@isi.edu

ABSTRACT

In educating emerging leaders to meet the challenges of tomorrow's non-traditional conflicts, the DoD must take advantage of new pedagogical and technological methods and venues that provide the learner with perceived risk reduction during education processes. The authors discuss how budding commanders must deeply and effectively experience geopolitical, historical, sociological and psychological material to improve their risk analyses and management to produce decisiveness in complex, diverse situations. An environment is described where they can engage regularly with lower thresholds for taking risks: emotional, intellectual, social and (virtual) physical. This will drive them to truly expand their "live" knowledge base. This paper sets out how High Performance Computing (HPC) is the catalytic enabler for creating complex innovative learning environments in which young leaders can most thoroughly engage with the dynamic situations that they must master to be most effective. The ability of HPC to manage manifold complex factors will allow the DoD to create learning modules that recognize and ameliorate the elements of risk-taking that the learner undergoes when faced with new knowledge. Didactic instruction should be almost entirely provided by this advance in computer-aided education, with the live instructor focusing on the role of coach and guide for the preparation before, and reflection after, the use of the virtual learning environment. There is a valuable cadre of highly experienced leadership instructors who are skilled in integrating didactic material with successful field experience. The DoD can develop the technology to leverage the capabilities of those few instructors to make their talents universally available by capturing their input for HPC-enabled virtual learning environments. The goal is to radically alter instructional interfaces to enhance vital pedagogical processes and thereby improve educational outcomes in fundamental and transformational ways. Documented support for the stated propositions and detailed analyses based on experience are set forth.

High-performance Computing Enables Simulations to Transform Education

Dan M. Davis

Information Sciences Institute
4676 Admiralty Way, Ste 1001, USC
Marina del Rey CA. 90292, U.S.A.

Thomas D. Gottschalk

Center for Advanced Computing Research
1200 E California MS158-79, Caltech
Pasadena CA 91125, U.S.A.

Laurel K. Davis

Next Generation Leaders, Inc.
Post Office Box 2573
Culver City, CA 90231. U.S.A

ABSTRACT

This paper presents the case that education in the 21st Century can only measure up to national needs if technologies developed in the simulation community, further enhanced by the power of high performance computing, are harnessed to supplant traditional didactic instruction. The authors cite their professional experiences in simulation, high performance computing and pedagogical studies to support their thesis that this implementation is not only required, it is feasible, supportable and affordable. Surveying and reporting on work in computer-aided education. This paper will discuss the pedagogical imperatives for group learning, risk management and “hero teacher” surrogates, all being optimally delivered with entity level simulations of varying types. Further, experience and research is adduced to support the thesis that effective implementation of this level of simulation is enabled only by, and is largely dependent upon, high performance computing, especially by the ready utility and acceptable costs of Linux clusters.

FPGA and GPU Accelerators for Linux Clusters

Dan M. Davis, James R. Moore and Robert F. Lucas

Information Sciences Institute, University of Southern California

4676 Admiralty Way, Suite 1001

Marina del Rey, CA 90292

{ddavis, jjmoore, rflucas} @isi.edu

Telephone: 310 822-1511

Abstract

The Test and Evaluation (T&E) community has been making great advances in using Modeling and Simulation (M&S) in their work. They would be even better served had they ready access to higher resolution, quicker turn-around times, more elements, and richer behavioral characteristics in their physics-based and entity-level simulations. As rapidly as it has been enabled to accomplish superior results, the T&E environment is still constrained by computational limits. High Performance Computing (HPC) can ameliorate those constraints. The use of Linux Clusters is one path to higher performance; the use of Field Programmable Gate Arrays (FPGAs) and Graphics Processing Units (GPU) as accelerators are two others. Merging these paths together holds even more promise. The authors report their experiences with the new HPCMP-provided 512 CPU (1024 core), GPU-enhanced Linux Cluster for the Joint Forces Command's Joint Experimentation Directorate (J9). They further relate their work on FPGAs as computational accelerators which bring with them the reprogrammable efficiency that are a complement to the GPUs powerful floating point efficacy. Basic concepts are laid out that underlie the use of FPGAs and GPUs as accelerators for intelligent agent, entity-level simulations and for multi-frontal attacks on sparse systems of linear equations. These two disparate fields will be used to show the broad range of capability improvements projected by the authors for FPGAs and GPUs. They discuss the use of the two accelerators in tandem as well. The simulation needs of the T&E community, the ability of FPGA- and GPU-enhanced clusters to respond to T&E needs, and the careful analysis of the intersection of these are explicitly discussed. Existing configurations and potential configurations of clusters are addressed and the potential increase in performance are identified and justified. Anticipated problems and solutions will all be reported objectively, as guides to the T&E community. The paths to reliable and timely capability enhancement will be fully explicated. Early characterization runs of a single CPU with GPU-enhanced extensions are reported.

Implementing a GPU-Enhanced Cluster for Large-Scale Simulations

Dan M. Davis & Gene Wagenbreth
Information Sciences Institute, Univ. of So. Calif.
Marina del Rey, California
[ddavis](mailto:ddavis@isi.edu) & [genew](mailto:genew@isi.edu) @isi.edu

ABSTRACT

The simulation community has often been hampered by constraints in computing: not enough resolution, not enough entities, not enough behavioral variants. Higher performance computers can ameliorate those constraints. The use of Linux Clusters is one path to higher performance; the use of Graphics Processing Units (GPU) as accelerators is another. Merging the two paths holds even more promise. The authors were two the principal architects of a successful proposal to the High Performance Computing Modernization Program (HPCMP) for a new 512 CPU (1024 core), GPU-enhanced Linux Cluster for the Joint Forces Command's Joint Experimentation Directorate (J9). In this paper, the basic theories underlying the use of GPUs as accelerators for intelligent agent, entity-level simulations are laid out, the previous research is surveyed and the ongoing efforts are outlined. The simulation needs of J9, the direction from HPCMP and the careful analysis of the intersection of these are explicitly discussed. The configuration of the cluster and the assumptions that led to the conclusion that GPUs might increase performance by a factor of two are carefully documented. The processes that led to that configuration, as delivered to JFCOM, will be specified and alternatives that were considered will be analyzed. Planning and implementation strategies are reviewed and justified. The paper will then report in detail about the execution of the actual installation and implementation of the JSAF simulation on the cluster. Issues, problems and solutions will all be reported objectively, as guides to the simulation community and as confirmation or rejection of early assumptions. Lessons learned and recommendations will be set out in detail. Original performance projections will be compared to actual benchmarking results using LINPACK and simulation performance. Early observed operational capabilities of interest will be proffered.

A GPU-Enhanced Cluster for Accelerated FMS

Dan M. Davis & Robert F. Lucas

Information Sciences Institute, Univ. of So. Calif.

Marina del Rey, California

[{ddavis & rflucas} @isi.edu](mailto:{ddavis & rflucas}@isi.edu)

ABSTRACT

The Forces Modeling and Simulation (FMS) community has often been hampered by constraints in computing: not enough resolution, not enough entities, not enough behavioral variants. Higher performance computers can ameliorate those constraints. The use of Linux Clusters is one path to higher performance; the use of Graphics Processing Units (GPU) as accelerators is another. Merging the two paths holds even more promise. There was successful proposal to the High Performance Computing Modernization Program (HPCMP) for a new 512 CPU (1024 core), GPU-enhanced Linux Cluster for the Joint Forces Command's Joint Experimentation Directorate (J9). The basic concepts underlying the use of GPUs as accelerators for intelligent agent, entity-level simulations are laid out. The simulation needs of J9, the direction from HPCMP and the careful analysis of the intersection of these are explicitly discussed. The configuration of the cluster and the assumptions that led to the conclusion that GPUs might increase performance by a factor of two are carefully documented. The paper will then report in detail about the execution of the actual installation and implementation of the JSAF simulation on the cluster. Issues, problems and solutions will all be reported objectively, as guides to the FMS community and as confirmation or rejection of early assumptions. Lessons learned and recommendations will be set out in detail. Early characterization runs of a single CPU with GPU-enhanced extensions will be reported.

Modeling Human Performance of Situation Awareness in Constructive Simulations

John J. Tran & Ke-Thia Yao
Information Sciences Institute
Marina del Rey, California
{jtran,kyao}@isi.edu

Philip Colon
Toyon Corporation
Goletta, California
philipc@toyon.com

Jacqueline M. Curiel & Michael D.
Alion Science and Technology
Alexandria, Virginia
{jcuriel,manhalt}@alionscience.com

ABSTRACT

Highly advanced sensor technologies give our military commanders a significant command and control (C2) advantage over our enemies during conflicts, particularly with respect to situation awareness (SA). The use of advanced sensor technology models in synthetic battlespace gives war fighters parallel advantages. Two accepted simulation methodologies for analyzing the impact of sensor technologies are through HITL experiments, such as Joint Urban Operations (JUO), which utilize sensor capabilities to assist human participants, and Monte Carlo constructive (MCC) simulations, which can be used to model human performance. In HITL experiments using Joint Semi-Automated Forces (JSAF), participants describe their SA using Situation Awareness Objects (SAOs, which then can be reconstructed using Endsley's (1995) three levels of SA (perception, comprehension, and prediction). MCC experiments, which are dominated by algorithmically determined behaviors, can be used to model SA. Sensor measurements currently can be fused to perceive individual entities, but do not have the capability to recognize groupings of entities, resulting only in partial perceptual SA. Furthermore, current sensor data fusion models do not produce the second and third levels of SA, comprehension and prediction.

This paper will report research efforts to utilize both methodologies to expand the use of SAOs beyond player declarations to the automatic generation of SAOs. We develop a method to organize events drawn from scenarios taken from HITL experiments using SAOs in order to develop situation awareness algorithms for the MCC runs. A comparison of these model-generated synthetic SAOs (SSAOs) to SAOs generated by human players can identify strengths and weakness in the SA models as well as identifying ways in which player performance can be improved.

Multifrontal GPGPU for MCAE

Robert F. Lucas, Gene Wagenbreth, John J. Tran, and Dan M. Davis
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, CA 90292

Email: {rflucas, genew, jtran & ddavis}@isi.edu

Abstract

Sparse systems of linear equations are computational bottlenecks in applications ranging from science to optimization. For many problems, including Mechanical Computer Aided Engineering (MCAE), iterative methods are unreliable and sparse matrix factorization is performed. Multifrontal sparse matrix factorization is often preferred and, by representing the sparse problem as a tree of dense systems, maps well to modern memory hierarchies. This allows effective use of BLAS-3 dense matrix arithmetic kernels. Graphics processing units (GPUs) are architected differently than general-purpose hosts and have an order-of-magnitude more single-precision floating point processing power. This paper explores the hypothesis that GPUs can accelerate the speed of a multifrontal linear solver, even when only processing a small number of the largest frontal matrices. We show that GPUs can more than double the throughput of the sparse matrix factorization. This in turn promises to offer a very cost-effective speedup to many problems in disciplines such as MCAE.

Sections: *Applications and Performance*

Keywords: *Computational solid mechanics and materials and application performance*

Implementing Multi-Abstraction Level Simulations: Enabling Consistency, Integration and Validation

Thomas D. Gottschalk

Center for Advanced Computing Research, Caltech
Pasadena, California
tdg@cacr.caltech.edu

Robert F. Lucas, Dan M. Davis

Information Sciences Institute, Univ. of So. Calif.
Marina del Rey, California
[rflucas](mailto:rflucas@isi.edu) or [ddavis](mailto:ddavis@isi.edu)@isi.edu

ABSTRACT

DoD requirements for training, analysis and evaluation require simulation technologies that provide realism and consistency across multiple abstraction layers. Abstraction dimensions range from resolution of entities (soldiers to battalions) to models of behavior paradigms (combat forces doctrines to social conduct predispositions). Both everyday users and General Officer commanders report and decry the lack of adequate interaction among the humans in the loop, the simulated forces and the social-urban interaction components. One example of this failure in abstraction consistency is the much reported aggregation/de-aggregation problem, which is regularly held to be intractable. Multiple resolutions are essential in addressing current simulation needs. A single simulation addressing all entities at all levels of resolution is simply not feasible, independent of available resources. The issue is one of synchronizing the component simulations, preventing the significant inconsistencies among different resolutions. The authors have advanced a new approach to overcome this obstacle and they are embarked upon research into this and other potential solutions that would have a significant impact across all of the services and all multi-abstraction simulations. The ultimate goal is the provision of “platform portable” technology to ensure realistic consistency between abstraction layers. Preliminary research is implementing proof of concept demonstrations *via* a simulation scenario, using a reduced set of parameters, driving an exemplar of forces simulation, the Corps Battlefield Simulator and a social modeling program, the Joint NonKinetic Effects Model. The authors lay out their view of the need, the problem, and the research plan. They discuss the choice of programs and compute platform for the experiments and present an overview of the architecture developed. Early results of the tests and implications of these results on integration and validation are advanced. They conclude by discussing future research requirements and architectural issues lying at the heart of more general, valid multiresolution simulation procedures.

A High Performance Route-Planning Technique for Dense Urban Simulations

**John J. Tran, Ke-Thia Yao, Gene Wagenbreth, Dan M. Davis and
Robert F. Lucas**

**Information Sciences Institute/USC
Marina del Rey, California
{jtran,kyao,genew,ddavis,rflucas}@isi.edu**

**David J. Bakeman
Nakuru Software Inc
Seattle, Washington
nakuru@comcast.net**

ABSTRACT

To exploit the explicit and implicit advantages of data parallelism and heavily threaded advanced multi-core processors, specifically the NVIDIA family of general purpose graphic processing units (GPGPU), research efforts such as "Accelerating Line of Sight Computation Using GPUs" (Manocha 2005) and "Implementing a GPU-Enhanced Cluster for Large-Scale Simulations" (Lucas 2007) addressed various problems found in military simulations. Yet other practical uses for the GPU in these types of simulation applications remain to be explored. An example application that has immediate use for a fast and large-scale graph-based construct is a route-planning algorithm found in complex urban conflict simulation, e.g. the Joint Semi-Automated Forces (JSAF) simulation. JSAF currently employs a heuristic A* search algorithm to do route planning for its millions of entities — the algorithm is sequential and thus very computationally expensive. Using the GPU, the JSAF simulation can off-load the route-planning component to the GPU and remove one of its major bottlenecks. The objective of this research effort is to build a framework that utilizes all the features and raw computational power of the GPU architecture to solve the above challenge. Our research effort addresses the many challenges of parallel programming on the GPU, e.g. data locality, massive thread counts, and race conditions, to name a few. Our project will greatly benefit the modeling and simulation community facing issues specific to route planning and of particular interest are those simulations dealing with dense urban environments, homeland security, and mass casualty and disaster simulations. We achieve this goal by providing a practical and seemingly "endless" source of raw computing powers found in GPUs for massively large graph-based family of problems.

Effectively Using a Large GPGPU-Enhanced Linux Cluster

Gene Wagenbreth, Robert F. Lucas and Dan M. Davis

Information Sciences Institute, Univ of So. Calif.

Marina del Rey, California

{genew, rflucas, ddavis}@isi.edu

ABSTRACT

Computing power units per dollar, per kilowatt and per square meter of computer floor footprint may be increased by using heterogeneous computing. The Joint Forces Command (JFCOM) has an urgent and continuing need for large-scale simulations. Both the training and the experimentation directorates must be able to effectively portray the battlespace of the future, often an urban setting with elaborate infrastructure, vast collateral damage possibilities and up to ten million civilians and vehicles. The compute power required is substantial. The authors report on their role in and work with the largest General Purpose Graphics Processing Unit (GPGPU)-enhanced Linux cluster of which they are aware: Joshua at JFCOM, which was awarded as a Dedicated High Performance Computing Project Investment (DHPI) project in 2007. Joshua's 256 nodes are enhanced with an NVIDIA 8800 GPU, each with two 2.33 GHz AMD dual-core Opterons and 16 GB of memory. The authors discuss the theoretical underpinnings that led them to propose such a computer, the process of acquiring it, its installation, early experience, and characterization. They then discuss the creation and their presentation of a course for users and programmers in the new Compute Unified Device Architecture (CUDA), and report on the success of this course. They will give a short precise of the course for those who may be inclined to seek out such an opportunity. They finally compare this programming model with several alternate programming models and compare the ease of programming GPGPUs with that of programming FPGAs and Cell processor chips. In this process the benchmarking and characterization approaches for several types of code are laid out and the results of the experiments are set forth. Several codes were considered, *e.g.* the traditional Linpack, the Multi-Frontal Sparse Matrix Solver, Route Planning algorithms, Line of Sight (LOS) and other agent-based simulation algorithms. The reasons for the final selection of codes for extensive characterization will be discussed. Performance data and optimization techniques used will be laid out in sufficient detail to assist others who are interested in the approach and assessing how effective it may be, if implemented in their environment. Future and expanded uses of the GPGPU acceleration technique and a description of logical programming candidates for this method are also considered in the conclusion section.

Multi-Abstraction Level Simulations: Enabling Consistency, Integration and Validation

Thomas D. Gottschalk

Center for Advanced Computing Research, Caltech
Pasadena, California
tdg@cacr.cacr.caltech.edu

Dan M. Davis

Information Sciences Institute, Univ of So. Calif.
Marina del Rey, California
ddavis@isi.edu

ABSTRACT

DoD requirements for training, analysis and evaluation require simulation technologies that provide realism and consistency across multiple abstraction layers. Abstraction dimensions range from resolution of entities (soldiers to battalions) to models of behavior paradigms (combat forces doctrines to social conduct predispositions). Both everyday users and General Officer commanders report and decry the lack of adequate interaction among the humans in the loop, the simulated forces and the social-urban interaction components. One example of this failure in abstraction consistency is the much reported aggregation/de-aggregation problem, which is regularly held to be intractable. Multiple resolutions are essential in addressing current simulation needs. A single simulation addressing all entities at all levels of resolution is simply not feasible, independent of available resources. The issue is one of synchronizing the component simulations, preventing the significant inconsistencies among different resolutions. The authors have advanced a new approach to overcome this obstacle and they are embarked upon research into this and other potential solutions that would have a significant impact across all of the services and all multi-abstraction simulations. The ultimate goal is the provision of “platform portable” technology to ensure realistic consistency between abstraction layers. Preliminary research is implementing proof of concept demonstrations via a simulation scenario, using a reduced set of parameters, driving an exemplar of forces simulation, the Corps Battlefield Simulator and a social modeling program, the Joint NonKinetic Effects Model. The authors lay out their view of the need, the problem, and the research plan. They discuss the choice of programs and compute platform for the experiments and present an overview of the architecture developed. Early results of the tests and implications of these results on integration and validation are advanced. They conclude by discussing future research requirements and architectural issues lying at the heart of more general, valid multi-resolution simulation procedures.

A High Performance Route-Planning Technique for Dense Urban Simulations

John J. Tran and Robert F. Lucas

Information Sciences Institute, Univ of So. Calif.

Marina del Rey, California

{jtran, rflucas}@isi.edu

ABSTRACT

To exploit the explicit and implicit advantages of data parallelism and heavily threaded modern multi-core processors, specifically the NVIDIA family of general purpose graphic processing unit (GPGPU), research efforts such "Accelerating Line of Sight Computation Using GPUs" [Manocha 2005] and "Implementing a GPU-Enhanced Cluster for Large-Scale Simulations" [Lucas 2007] addressed the various problems found in military simulations,. Yet there remain many other practical uses for the GPU in these types of simulation applications. An example application that has immediate use for a fast and large-scale graph-based construct is a route-planning algorithm found in complex urban conflict simulation, e.g. the Joint Semi-Automated Forces (JSAF) simulation. JSAF currently employs a heuristic A* search algorithm to do route planning for its millions of entities -- the algorithm is sequential and thus very computationally expensive. Using the GPU, the JSAF simulation can off-load the route planning component to the GPU and remove one of its major bottlenecks.

The objective of this research effort is to build a framework that utilizes all the features and raw computational power of the GPU architecture to solve the above challenge. Our research effort addresses the many challenges of parallel programming on the GPU, e.g. data locality, massive thread counts, and race conditions, to name a few. Our project will greatly benefit the modeling and simulation community facing issues specific to route-planning and of particular interest are those simulations dealing with dense urban environments, homeland security, and mass casualty and disaster simulations. We achieve this goal by providing a practical and seemingly "endless" source of raw computing powers found in GPUs for massively large graph-based family of problems.

Data Analysis for Massively Distributed Simulations

Ke-Thia Yao, Robert F. Lucas, Craig E. Ward, Gene

USC Information Sciences Institute
Marina del Rey, California
{kyao, rflucas, cward, genew}@isi.edu

Thomas D. Gottschalk

California Institute of Technol-
Pasadena, California
tdg@cacr.caltech.edu

ABSTRACT

More computing power allows increases in the fidelity of simulations. Fast networking allows large clusters of high performance computing resources, often distributed across wide geographic areas, to be brought to bear on the simulations. This increase in fidelity has correspondingly increased the volumes of data simulations are capable of generating. Coordinating distant computing resources and making sense of this mass of data is a problem that must be addressed. Unless data are analyzed and converted into information, simulations will provide no useful knowledge. This paper reports on experiments using distributed analysis, particularly the Apache Hadoop framework, to address the analysis issues and suggests directions for enhancing the analysis capabilities to keep pace with the data generating capabilities found in modern simulation environments. Hadoop provides a scalable, but conceptually simple, distributed computation paradigm based on map/reduce operations implemented over a highly parallel, distributed filesystem. We developed map/reduce implementations of K-Means and Expectation-Maximization data mining algorithms that take advantage of the Hadoop framework. The Hadoop filesystem dramatically improves the disk scan time needed by these iterative data mining algorithms. We ran these algorithms across multiple Linux clusters over specially reserved high speed networks. The results of these experiments point to potential enhancements for Hadoop and other analysis tools.

FLOPS per Watt: Heterogeneous-Computing's Approach to DoD Imperatives

**Dan M. Davis, Robert F. Lucas, Gene Wagenbreth,
John J. Tran & Joel Agaloff**

Information Sciences Institute/Univ. of So. Calif.
Marina del Rey, California
{ ddavis, rflucas, genew, jtran, agaloff }@isi.edu

Thomas D. Gottschalk
Center for Advanced Computing Research

California Institute of Technology
Pasadena, California
tdg@cacr.caltech.edu

ABSTRACT

Electrical power used in computing is increasingly a vital factor in all computing, from laptops to PetaFLOPS. Cost, portability, ecological concerns and hardware life are all negatively impacted by burgeoning power requirements. More than the rest of the world, the U.S. DoD has special requirements to restrain the use of electrical power, ranging from battery life for devices in the field to environmental responsibility for major DoD Supercomputing Centers. The authors will discuss the special insights they have gained into the implementation of one technique, the use of General Purpose Graphics Processing Units as heterogeneous processors and they will further outline the state of the art in the field of power reduction techniques, ranging from IBM's Blue Gene series to Prof. William Dally's Efficient Low-power Microprocessor (ELM) approach and compare and contrast them with the experience of the authors on JFCOM's Joshua, a 256 node, GPGPU enhanced cluster. Using GPGPUs to effectively handle computationally intensive activity "spikes" is manifestly germane to defense computational needs. Quantitatively, the authors will report on three specific aspects their use of GPGPUs: programming environment constraints and opportunities, performance of codes modified in several areas of computational science and the FLOPS per Watt parameter in a wide range of software and hardware configurations. An overview of algorithmic design and implementation strategies will be laid out. Actual working code segments will be discussed and explained, along with the design rationale behind them. The authors' experience in training other DoD users in this technique will assist program managers in scoping training requirements. This data should allow other DoD researchers and users to effectively anticipate the benefits of this approach as far as their own code is concerned and further, it should enable them to effectively evaluate the varying benefits of all of the approaches currently extant.

A Framework for Team Situation Awareness in Synthetic Battlespace

*** Primary Author ***

Dr. Jacqueline Curiel

jcuriel@AlionScience.com

Biography:

Jacqueline M. Curiel is a research psychologist at Alion Science and Technology and a co-founder of Behavioral Cognition LLC, a consulting company specializing in behavioral research. She received both her M. A. and PhD degrees in Psychology from the University of Notre Dame, where her research focused on spatial cognition and mental representations in narrative comprehension.

*** Secondary Author ***

Captain John Tran

jtran@isi.edu

Situation awareness (SA) in group contexts such as in human-in-the-loop (HITL) experiments can differ markedly from other military contexts where performance centers on the individual (e.g., fighter pilot SA). One obvious difference is that in group contexts, information relevant to the situation is obtained and used by more than one individual. As a result, HITL players bring to gameplay backgrounds that vary in terms of level of experience, skills/ abilities, and prior knowledge and so contribute differentially to the information-gathering and sense making processes involved in SA. While most definitions of SA have focused on the internal representations and processes of the individual other attempts have distinguished between the individual and group both in terms of the unit of analysis (individual vs. system) and in identifying the processes and representations involving Team SA (e.g., distribution of information within the system and the dynamic coordination of this information across time). This latter focus involves distributed cognition.

In this paper we develop a framework for situation awareness within the context of synthetic battlespace that incorporates ideas about individual and Team SA to assess the contribution of individual players, the distributed cognitive system, and the performance of the team as a whole using objective and subjective measures of evaluation. Using data from a HITL experiment we will illustrate concepts relevant to this framework. It is our intent that this framework generalizes to other dynamic group contexts. Among the advantages of this approach are that it increases opportunities for learning by separating out individual performance and that it provides a guide for developing more effective training software and techniques, both of which will ultimately contribute to an increase in mission effectiveness.

Experimental Use of a Large GPGPU-Enhanced Linux Cluster

Robert F. Lucas, Dan M. Davis, Gene Wagenbreth, and John J. Tran

Information Sciences Institute, University of Southern California, Marina del Rey, California

{rflucas, ddavis, genew, and jtran} @isi.edu

and

Andy Ceranowicz, Alion Science and Technology

ACeranowicz@alionscience.com

Introduction This year the Joint Forces Command (JFCOM) put the General Purpose Graphics Processing Units (GPGPUs) enhanced cluster, Joshua, into production use. This cluster has demonstrable advantages in the optimization of performance for a large number of algorithms, hence should also be useful in other cluster configurations. Such a Linux cluster was seen as interesting to the HPCMP and one was provided to the Joint Forces Command in Suffolk Virginia, as that they had a manifest need for such computing power. Having worked with the new JFCOM GPU-enhanced Linux Cluster, the authors relate their experiences, lessons-learned and insights. They report the porting of several code modules to effectively use the GPUs and the use of the cluster to simulate ten million “CultureSim” agent-based entities.

Objective The ultimate objective of this research is to provide JFCOM with an order of magnitude power of scale in computing in the demanding research environment at JFCOM. This enabled them to continue to develop, explore, test, and validate 21st century battlespace concepts. The specific goal is to enhance global-scale, computer-generated experimentation by sustaining more than 2,000,000 entities on appropriate terrain with valid phenomenology. That goal was exceeded. The authors report they still are confident that there will eventually be an order of magnitude increase in the stated goal.

Methodology The method employed was to use existing DoD simulation codes on the advanced Linux clusters operated by JFCOM. The improved cluster reported herein supplants the original JFCOM J9 DC clusters with new upgraded 64-bit CPUs and enhanced with nVidia 8800 GPUs. Further, the authors have begun to modify legacy codes to make optimal use of the GPUs’ substantial processing power. Initially, the major driver for the FMS community’s use of accelerator-enhanced nodes was the need for faster processing to accomplish line-of-sight calculations. The first experiments were used as a training evolution on a smaller code set, one also amenable to GPU acceleration, to facilitate the programming and hasten the experimentation insights.

Results The learning curve for the use of the new C-like CUDA code for GPU non-graphics processing was found to be manageable. It was demonstrated that the GPU could be very effective at reducing the time spent factoring the large frontal matrices near the root of the elimination tree in the strategic calculation approach. The GPU accelerated the overall factorization at close to the factor of two originally hypothesized. One result that has already been achieved is that the goal of 2M entities was exceeded by a factor of five during a 10,000,000 entity run. This was still not seen as a maximum, as no hard barriers were observed, so further growth is anticipated, perhaps in entity complexity rather than number.

GPGPUs as Heterogeneous Accelerators in a 256 Node Linux Cluster

Robert F. Lucas, Gene Wagenbreth, John J. Tran, Dan M. Davis and Ke-Thia Yao

Information Sciences Institute/USC
Marina del Rey, California
{ rflucas, genew, jtran, ddavis, kyao }@isi.edu

ABSTRACT

It is asserted that units of computational power per dollar, per kilowatt and per square meter of computer room floor footprint may be increased by using heterogeneous computing. Simulation users must be able to effectively portray populated environments of the future, often in urban settings with elaborate infrastructure, complex emergency response possibilities and appropriate populations, up to ten million civilians and vehicles. The compute power required is substantial. The authors report on their role in and work with the largest General Purpose Graphics Processing Unit (GPGPU)-enhanced Linux cluster of which they are aware: Joshua at JFCOM. Each of Joshua's 256 nodes are enhanced with an NVIDIA 8800 GPU, each with two 2.33 GHz AMD dual-core Opterons and 16 GB of memory. The authors discuss the theoretical underpinnings that led them to propose such a computer, the process of acquiring it, its installation, early experience, and characterization. They then discuss the creation and their presentation of courses for users and programmers in the new Compute Unified Device Architecture (CUDA), and report on the success of these courses. They finally compare this programming model with several alternate programming models and compare the ease of programming GPGPUs with that of programming FPGAs and Cell processor chips. In this process the benchmarking and characterization approaches for several types of code are laid out and the results of the experiments is set forth. Several codes were considered, *e.g.* the traditional Linpack, the Multi-Frontal Sparse Matrix Solver, Route Planning algorithms, Line of Sight (LOS) and other agent-based simulation algorithms. They discuss attempts to exploit the explicit and implicit advantages of data parallelism and heavily threaded advanced multi-core processors, specifically the NVIDIA family of general purpose graphic processing units (GPGPU). The reasons for the final selection of codes for extensive characterization will be discussed. Future and expanded uses of the GPGPU acceleration technique and a description of logical programming candidates for this method are also considered in the conclusion section. An example application that has immediate use for a fast and large-scale graph-based construct is a route-planning algorithm found in complex urban conflict simulation, *e.g.* the Joint Semi-Automated Forces (JSAF) simulation. JSAF currently employs a heuristic A* search algorithm to do route planning for its millions of entities — the algorithm is sequential and thus very computationally expensive. Using the GPU, the JSAF simulation can off-load the route-planning component to the GPU and remove one of its major bottlenecks. Our research effort addresses the many challenges of parallel programming on the GPU, *e.g.* data locality, massive thread counts, and race conditions.

APPLYING HPCC TECHNOLOGY TO SYSTEMIC DIVERSITY DILEMMAS

Dan M. Davis	Andrea M. Baer	Laurel K. Davis	Thomas D. Gottschalk
<i>Information Sciences Institute</i>	<i>National Portrait Gallery</i>	<i>Leadership Training Group</i>	<i>Cntr. for Advanced Computing Resch.</i>
<i>University of Southern California</i>	<i>Smithsonian Institution</i>	<i>Next Generation Leaders</i>	<i>California Institute of Technology</i>

The authors present their decades of research and development experience in applying High Performance Computing and Communications (HPCC) technology to defense simulations and training. They discuss current dilemmas in education, including increasingly diverse classrooms and identify those for which HPCC technologies may hold the answer. Further they adduce evidence to support their thesis that such technology could be applied, resulting in attractive cost/benefit ratios, increased pedagogical efficacy, fewer teacher administrative burdens and, most importantly, more effective responses to diversity-related needs found in several disparate dimensions. They recount their hands-on experience in pre-college education environments, their compilation of data on classroom teacher perceptions and their justification and procurement of HPCC assets to meet otherwise daunting challenges. A special feature of this work is its concentration on teacher-centered services that are relevant, accessible and controllable by less technically sophisticated teachers, especially in early education environments. Rather than imposing that which is technically exciting, they focus on what teachers and learners want and need. Personalizing individual instruction, to both enable each student to learn and to address the identified classroom dilemmas, can arguably be best served by well-designed HPCC-supported platforms and modules. The extensibility of this approach to informal education is explored in the context of museum education.